

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION  
International Bureau

## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 7 : C12N 15/12, C07K 14/51, 14/495, C12N 15/63, 5/10, C07K 16/22, C12Q 1/68, C12N 15/62, A61K 38/18, A61P 19/10, G01N 33/53, A01K 67/027		A1	(11) International Publication Number: <b>WO 00/32773</b>
(21) International Application Number: PCT/US99/27990		(43) International Publication Date: 8 June 2000 (08.06.00)	
(22) International Filing Date: 24 November 1999 (24.11.99)		(74) Agent: MCMASTERS, David, D.; Seed and Berry LLP, Suite 6300, 701 Fifth Avenue, Seattle, WA 98104-7092 (US).	
(30) Priority Data: 60/110,283 27 November 1998 (27.11.98) US		(81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).	
(71) Applicant (for all designated States except US): DARWIN DISCOVERY LTD. [GB/GB]; Cambridge Science Park, Milton Road, Cambridge, Cambridgeshire CB4 4WE (GB).		Published With international search report. Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.	
(72) Inventors; and (75) Inventors/Applicants (for US only): BRUNKOW, Mary, E. [US/US]; 9829 Triton Drive NW, Seattle, WA 98111 (US). GALAS, David, J. [US/US]; 854 Guanajuato Drive, Claremont, CA 91711 (US). KOVACEVICH, Brian [US/US]; 4308 N.E. 6th Place, Renton, WA 98059 (US). MULLIGAN, John, T. [US/US]; 5823 17th Avenue Northeast, Seattle, WA 98105 (US). PAEPER, Bryan, W. [US/US]; 1617 Summit Avenue #43, Seattle, WA 98122 (US). VAN NESS, Jeffrey [US/US]; 10020 49th Avenue Northeast, Seattle, WA 98125 (US). WINKLER, David, G. [US/US]; 7037 20th Avenue NE, Seattle, WA 98115 (US).			

(54) Title: COMPOSITIONS AND METHODS FOR INCREASING BONE MINERALIZATION

## (57) Abstract

A novel class or family of TGF- $\beta$  binding proteins is disclosed. Also disclosed are assays for selecting molecules for increasing bone mineralization and methods for utilizing such molecules.

## Common Cysteine Backbone

1	human_gremlin.pro	-----	50
	human_cerberus.pro	MHLLLFQLLV LPLGKTRH QGQRONQSSL SPVLLPRNQR ELPTGNHEEA	
	human_dan.pro	-----	
	human_beer.pro	-----	
51	human_gremlin.pro	-----M SRTAYTVOAL LLLLOTLLPA AEGKKKQSQG	100
	human_cerberus.pro	BEKPDLPVAV PHLVAT SPA GEGQRQREKM LSRPGRFWKK PEREMHPSRD	
	human_dan.pro	-----	
	human_beer.pro	-----MQLPLA LCLVCLLVHT	
101	human_gremlin.pro	AI.PPPDKAQ HNDSEOTQSP QOQSRNROR QGQRQTAMPO EGVLESSQEA	150
	human_cerberus.pro	SDSEPPFPPT QSLIQPID G MOMEKSPLE EAKKPMWHFM FRKTPASQGV	
	human_dan.pro	-----MLRVLVGAVL PAMLLAAPPF	
	human_beer.pro	AFRVVEGGQW QAFKNDATETI IPELQEYPEP PPELENNKTH NRAENQGRPF	
151	human_gremlin.pro	LHVTERKYLK RDWCKTQPLK QTIHEEGCNS RTIINRF.CY GQCNSFYTPR	200
	human_cerberus.pro	ILPIKSHEVH WETCRTVPPS QTITHEGCEK VVVQNNL.CF GKCSVHFP.	
	human_dan.pro	INKLALFPDK SAWCEAKNIT QIVGHSQCEA KSIQNR.A.CL GQCFSYSVFN	
	human_beer.pro	HHPPFETKQVS EYSCRELHFT RYVTDGCPERS AKPVELTVCS GQCQPARLLP	
201	human_gremlin.pro	HIRKSEGSFO SCSP...CKP KKTFTMMVTL NCPQLQPTTK K.KRVTRVKQ	250
	human_cerberus.pro	..GAAQHSHT SCSH...CLP AKPTTNHLPL NCTELSSVVK V...VHLVEE	
	human_dan.pro	TFPQSTESLV HCDS...CMP AQSMNEIVTL ECPQHEEVPR VDKLVEKILH	
	human_beer.pro	NAIGRGKWHR PSQPDPRCIP DRYRAQRVOL LCPQGSAPRA RKVRLVAS..	
251	human_gremlin.pro	CRC.ISIDLQ -----	300
	human_cerberus.pro	CQCKVTEHE DGHILHAGSO DSFIPQVSA -----	
	human_dan.pro	CSCQACQKEP SHEGLSVVYO GEDGPGSOPQ THPHPHPHPH PGQQTPEPED	
	human_beer.pro	CKCKRLTRFH NQSELKDFOT EAARPQKORK PRPRARSACA NQASELENAY-	
301	human_gremlin.pro	-----	314
	human_cerberus.pro	-----	
	human_dan.pro	PPGAPHTEEE GAED -----	
	human_beer.pro	-----	

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Larvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

## COMPOSITIONS AND METHODS FOR INCREASING BONE MINERALIZATION

### TECHNICAL FIELD

The present invention relates generally to pharmaceutical products and methods and, more specifically, to methods and compositions suitable for increasing the mineral content of bone. Such compositions and methods may be utilized to treat a wide variety of conditions, including for example, osteopenia, osteoporosis, fractures and other disorders in which low bone mineral density are a hallmark of the disease.

### BACKGROUND OF THE INVENTION

Two or three distinct phases of changes to bone mass occur over the life of an individual (see Riggs, *West J. Med.* 154:63-77, 1991). The first phase occurs in both men and women, and proceeds to attainment of a peak bone mass. This first phase is achieved through linear growth of the endochondral growth plates, and radial growth due to a rate of periosteal apposition. The second phase begins around age 30 for trabecular bone (flat bones such as the vertebrae and pelvis) and about age 40 for cortical bone (e.g., long bones found in the limbs) and continues to old age. This phase is characterized by slow bone loss, and occurs in both men and women. In women, a third phase of bone loss also occurs, most likely due to postmenopausal estrogen deficiencies. During this phase alone, women may lose an additional 10% of bone mass from the cortical bone and 25% from the trabecular compartment (see Riggs, *supra*).

Loss of bone mineral content can be caused by a wide variety of conditions, and may result in significant medical problems. For example, osteoporosis is a debilitating disease in humans characterized by marked decreases in skeletal bone mass and mineral density, structural deterioration of bone including degradation of bone microarchitecture and corresponding increases in bone fragility and susceptibility to fracture in afflicted individuals. Osteoporosis in humans is preceded by clinical osteopenia (bone mineral density that is greater than one standard deviation but less than 2.5 standard deviations below the mean value for young adult bone), a condition found in approximately 25 million people in the United States. Another 7-8 million patients in the United States have been diagnosed with clinical osteoporosis (defined as bone mineral content greater than 2.5 standard deviations below that of mature young adult bone). Osteoporosis is one of the most expensive diseases for the health care

system, costing tens of billions of dollars annually in the United States. In addition to health care-related costs, long-term residential care and lost working days add to the financial and social costs of this disease. Worldwide approximately 75 million people are at risk for osteoporosis.

5           The frequency of osteoporosis in the human population increases with age, and among Caucasians is predominant in women (who comprise 80% of the osteoporosis patient pool in the United States). The increased fragility and susceptibility to fracture of skeletal bone in the aged is aggravated by the greater risk of  
10       fractures are reported in the United States each year. Fractured hips, wrists, and vertebrae are among the most common injuries associated with osteoporosis. Hip fractures in particular are extremely uncomfortable and expensive for the patient, and for women correlate with high rates of mortality and morbidity.

          Although osteoporosis has been defined as an increase in the risk of  
15       fracture due to decreased bone mass, none of the presently available treatments for skeletal disorders can substantially increase the bone density of adults. There is a strong perception among all physicians that drugs are needed which could increase bone density in adults, particularly in the bones of the wrist, spinal column and hip that are at risk in osteopenia and osteoporosis.

20           Current strategies for the prevention of osteoporosis may offer some benefit to individuals but cannot ensure resolution of the disease. These strategies include moderating physical activity (particularly in weight-bearing activities) with the onset of advanced age, including adequate calcium in the diet, and avoiding consumption of products containing alcohol or tobacco. For patients presenting with  
25       clinical osteopenia or osteoporosis, all current therapeutic drugs and strategies are directed to reducing further loss of bone mass by inhibiting the process of bone absorption, a natural component of the bone remodeling process that occurs constitutively.

          For example, estrogen is now being prescribed to retard bone loss.  
30       There is, however, some controversy over whether there is any long term benefit to patients and whether there is any effect at all on patients over 75 years old. Moreover, use of estrogen is believed to increase the risk of breast and endometrial cancer.

          High doses of dietary calcium, with or without vitamin D has also been suggested for postmenopausal women. However, high doses of calcium can often have  
35       unpleasant gastrointestinal side effects, and serum and urinary calcium levels must be continuously monitored (see Khosla and Riggs, *Mayo Clin. Proc.* 70:978-982, 1995).



Other therapeutics which have been suggested include calcitonin, bisphosphonates, anabolic steroids and sodium fluoride. Such therapeutics however, have undesirable side effects (e.g., calcitonin and steroids may cause nausea and provoke an immune reaction, bisphosphonates and sodium fluoride may inhibit repair of fractures, even though bone density increases modestly) that may prevent their usage (see Khosla and Riggs, *supra*).

No currently practiced therapeutic strategy involves a drug that stimulates or enhances the growth of new bone mass. The present invention provides compositions and methods which can be utilized to increase bone mineralization, and thus may be utilized to treat a wide variety of conditions where it is desired to increase bone mass. Further, the present invention provides other, related advantages.

## SUMMARY OF THE INVENTION

As noted above, the present invention provides a novel class or family of TGF-beta binding-proteins, as well as assays for selecting compounds which increase bone mineral content and bone mineral density, compounds which increase bone mineral content and bone mineral density and methods for utilizing such compounds in the treatment or prevention of a wide variety of conditions.

Within one aspect of the present invention, isolated nucleic acid molecules are provided, wherein said nucleic acid molecules are selected from the group consisting of: (a) an isolated nucleic acid molecule comprising sequence ID Nos. 1, 5, 7, 9, 11, 13, or, 15, or complementary sequence thereof; (b) an isolated nucleic acid molecule that specifically hybridizes to the nucleic acid molecule of (a) under conditions of high stringency; and (c) an isolated nucleic acid that encodes a TGF-beta binding-protein according to (a) or (b). Within related aspects of the present invention, isolated nucleic acid molecules are provided based upon hybridization to only a portion of one of the above-identified sequences (e.g., for (a) hybridization may be to a probe of at least 20, 25, 50, or 100 nucleotides selected from nucleotides 156 to 539 or 555 to 687 of Sequence ID No. 1). As should be readily evident, the necessary stringency to be utilized for hybridization may vary based upon the size of the probe. For example, for a 25-mer probe high stringency conditions could include: 60 mM Tris pH 8.0, 2 mM EDTA, 5x Denhardt's, 6x SSC, 0.1% (w/v) N-laurylsarcosine, 0.5% (w/v) NP-40 (nonidet P-40) overnight at 45 degrees C, followed by two washes with 0.2x SSC / 0.1% SDS at 45-50 degrees. For a 100-mer probe under low stringency conditions, suitable conditions might include the following: 5x SSPE, 5x Denhardt's, and 0.5% SDS overnight at 42-50 degrees, followed by two washes with 2x SSPE (or 2x SSC)

/0.1% SDS at 42-50 degrees.

Within related aspects of the present invention, isolated nucleic acid molecules are provided which have homology to Sequence ID Nos. 1, 5, 7, 9, 11, 13, or 15, at a 50%, 60%, 75%, 80%, 90%, 95%, or 98% level of homology utilizing a Wilbur-Lipman algorithm. Representative examples of such isolated molecules include, for example, nucleic acid molecules which encode a protein comprising Sequence ID NOs. 2, 6, 10, 12, 14, or 16, or have homology to these sequences at a level of 50%, 60%, 75%, 80%, 90%, 95%, or 98% level of homology utilizing a Lipman-Pearson algorithm.

Isolated nucleic acid molecules are typically less than 100kb in size, and, within certain embodiments, less than 50kb, 25kb, 10kb, or even 5kb in size. Further, isolated nucleic acid molecules, within other embodiments, do not exist in a "library" of other unrelated nucleic acid molecules (*e.g.*, a subclone BAC such as described in GenBank Accession No. AC003098 and EMB No. AQ171546). However, isolated nucleic acid molecules can be found in libraries of related molecules (*e.g.*, for shuffling, such as is described in U.S. Patent Nos. 5,837,458; 5,830,721; and 5,811,238). Finally, isolated nucleic acid molecules as described herein do not include nucleic acid molecules which encode Dan, Cerberus, Gremlin, or SCGF (U.S. Patent No. 5,780,263).

Also provided by the present invention are cloning vectors which contain the above-noted nucleic acid molecules, and expression vectors which comprise a promoter (*e.g.*, a regulatory sequence) operably linked to one of the above-noted nucleic acid molecules. Representative examples of suitable promoters include tissue-specific promoters, and viral - based promoters (*e.g.*, CMV-based promoters such as CMV I-E, SV40 early promoter, and MuLV LTR). Expression vectors may also be based upon, or derived from viruses (*e.g.*, a "viral vector"). Representative examples of viral vectors include herpes simplex viral vectors, adenoviral vectors, adenovirus-associated viral vectors and retroviral vectors. Also provided are host cells containing or comprising any of above-noted vectors (including for example, host cells of human, monkey, dog, rat, or mouse origin).

Within other aspects of the present invention, methods of producing TGF-beta binding-proteins are provided, comprising the step of culturing the aforementioned host cell containing vector under conditions and for a time sufficient to produce the TGF-beta binding protein. Within further embodiments, the protein produced by this method may be further purified (*e.g.*, by column chromatography, affinity purification, and the like). Hence, isolated proteins which are encoded by the

above-noted nucleic acid molecules (*e.g.*, Sequence ID NOs. 2, 4, 6, 8, 10, 12, 14, or 16) may be readily produced given the disclosure of the subject application.

It should also be noted that the aforementioned proteins, or fragments thereof, may be produced as fusion proteins. For example, within one aspect fusion proteins are provided comprising a first polypeptide segment comprising a TGF-beta binding-protein encoded by a nucleic acid molecule as described above, or a portion thereof of at least 10, 20, 30, 50, or 100 amino acids in length, and a second polypeptide segment comprising a non-TGF-beta binding-protein. Within certain embodiments, the second polypeptide may be a tag suitable for purification or recognition (*e.g.*, a polypeptide comprising multiple anionic amino acid residues – see U.S. Patent No. 4,851,341), a marker (*e.g.*, green fluorescent protein, or alkaline phosphatase), or a toxic molecule (*e.g.*, ricin).

Within another aspect of the present invention, antibodies are provided which are capable of specifically binding the above-described class of TGF-beta binding proteins (*e.g.*, human BEER). Within various embodiments, the antibody may be a polyclonal antibody, or a monoclonal antibody (*e.g.*, of human or murine origin). Within further embodiments, the antibody is a fragment of an antibody which retains the binding characteristics of a whole antibody (*e.g.*, an F(ab')<sub>2</sub>, F(ab)<sub>2</sub>, Fab', Fab, or Fv fragment, or even a CDR). Also provided are hybridomas and other cells which are capable of producing or expressing the aforementioned antibodies.

Within related aspects of the invention, methods are provided detecting a TGF-beta binding protein, comprising the steps of incubating an antibody as described above under conditions and for a time sufficient to permit said antibody to bind to a TGF-beta binding protein, and detecting the binding. Within various embodiments the antibody may be bound to a solid support to facilitate washing or separation, and/or labeled (*e.g.*, with a marker selected from the group consisting of enzymes, fluorescent proteins, and radioisotopes).

Within other aspects of the present invention, isolated oligonucleotides are provided which hybridize to a nucleic acid molecule according to Sequence ID NOs. 1, 3, 5, 7, 9, 11, 13, 15, 17, or 18 or the complement thereto, under conditions of high stringency. Within further embodiments, the oligonucleotide may be found in the sequence which encodes Sequence ID Nos. 2, 4, 6, 8, 10, 12, 14, or 16. Within certain embodiments, the oligonucleotide is at least 15, 20, 30, 50, or 100 nucleotides in length. Within further embodiments, the oligonucleotide is labeled with another molecule (*e.g.*, an enzyme, fluorescent molecule, or radioisotope). Also provided are primers which are capable of specifically amplifying all or a portion of the above-

mentioned nucleic acid molecules which encode TGF-beta binding-proteins. As utilized herein, the term "specifically amplifying" should be understood to refer to primers which amplify the aforementioned TGF-beta binding-proteins, and not other TGF-beta binding proteins such as Dan, Cerberus, Gremlin, or SCGF (U.S. Patent No. 5,780,263).

Within related aspects of the present invention, methods are provided for detecting a nucleic acid molecule which encodes a TGF-beta binding protein, comprising the steps of incubating an oligonucleotide as described above under conditions of high stringency, and detecting hybridization of said oligonucleotide. Within certain embodiments, the oligonucleotide may be labeled and/or bound to a solid support.

Within other aspects of the present invention, ribozymes are provided which are capable of cleaving RNA which encodes one of the above-mentioned TGF-beta binding-proteins (e.g., Sequence ID NOs. 2, 6, 8, 10, 12, 14, or 16). Such ribozymes may be composed of DNA, RNA (including 2'-O-methyl ribonucleic acids), nucleic acid analogs (e.g., nucleic acids having phosphorothioate linkages) or mixtures thereof. Also provided are nucleic acid molecules (e.g., DNA or cDNA) which encode these ribozymes, and vectors which are capable of expressing or producing the ribozymes. Representative examples of vectors include plasmids, retrotransposons, cosmids, and viral-based vectors (e.g., viral vectors generated at least in part from a retrovirus, adenovirus, or, adeno-associated virus). Also provided are host cells (e.g., human, dog, rat, or mouse cells) which contain these vectors. In certain embodiments, the host cell may be stably transformed with the vector.

Within further aspects of the invention, methods are provided for producing ribozymes either synthetically, or by *in vitro* or *in vivo* transcription. Within further embodiments, the ribozymes so produced may be further purified and/or formulated into pharmaceutical compositions (e.g., the ribozyme or nucleic acid molecule encoding the ribozyme along with a pharmaceutically acceptable carrier or diluent). Similarly, the antisense oligonucleotides and antibodies or other selected molecules described herein may be formulated into pharmaceutical compositions.

Within other aspects of the present invention, antisense oligonucleotides are provided comprising a nucleic acid molecule which hybridizes to a nucleic acid molecule according to Sequence ID NOs. 1, 3, 5, 7, 9, 11, 13, or 15, or the complement thereto, and wherein said oligonucleotide inhibits the expression of TGF-beta binding protein as described herein (e.g., human BEER). Within various embodiments, the oligonucleotide is 15, 20, 25, 30, 35, 40, or 50 nucleotides in length. Preferably, the

oligonucleotide is less than 100, 75, or 60 nucleotides in length. As should be readily evident, the oligonucleotide may be comprised of one or more nucleic acid analogs, ribonucleic acids, or deoxyribonucleic acids. Further, the oligonucleotide may be modified by one or more linkages, including for example, covalent linkage such as a phosphorothioate linkage, a phosphotriester linkage, a methyl phosphonate linkage, a methylene(methylimino) linkage, a morpholino linkage, an amide linkage, a polyamide linkage, a short chain alkyl intersugar linkage, a cycloalkyl intersugar linkage, a short chain heteroatomic intersugar linkage and a heterocyclic intersugar linkage. One representative example of a chimeric oligonucleotide is provided in U.S. Patent No. 5,989,912.

Within yet another aspect of the present invention, methods are provided for increasing bone mineralization, comprising introducing into a warm-blooded animal an effective amount of the ribozyme as described above. Within related aspects, such methods comprise the step of introducing into a patient an effective amount of the nucleic acid molecule or vector as described herein which is capable of producing the desired ribozyme, under conditions favoring transcription of the nucleic acid molecule to produce the ribozyme.

Within other aspects of the invention transgenic, non-human animals are provided. Within one embodiment a transgenic animal is provided whose germ cells and somatic cells contain a nucleic acid molecule encoding a TGF-beta binding-protein as described above which is operably linked to a promoter effective for the expression of the gene, the gene being introduced into the animal, or an ancestor of the animal, at an embryonic stage, with the proviso that said animal is not a human. Within other embodiments, transgenic knockout animals are provided, comprising an animal whose germ cells and somatic cells comprise a disruption of at least one allele of an endogenous nucleic acid molecule which hybridizes to a nucleic acid molecule which encodes a TGF-binding protein as described herein, wherein the disruption prevents transcription of messenger RNA from said allele as compared to an animal without the disruption, with the proviso that the animal is not a human. Within various embodiments, the disruption is a nucleic acid deletion, substitution, or, insertion. Within other embodiments the transgenic animal is a mouse, rat, sheep, pig, or dog.

Within further aspects of the invention, kits are provided for the detection of TGF-beta binding-protein gene expression, comprising a container that comprises a nucleic acid molecule, wherein the nucleic acid molecule is selected from the group consisting of (a) a nucleic acid molecule comprising the nucleotide sequence of SEQ ID NOs: 1, 3, 5, 7, 9, 11, 13, or 15; (b) a nucleic acid molecule comprising the

complement of the nucleotide sequence of (a); (c) a nucleic acid molecule that is a fragment of (a) or (b) of at least 15, 20 30, 50, 75, or, 100 nucleotides in length. Also provided are kits for the detection of a TGF-beta binding-protein which comprise a container that comprise one of the TGF-beta binding protein antibodies described  
5 herein.

For example, within one aspect of the present invention methods are provided for determining whether a selected molecule is capable of increasing bone mineral content, comprising the steps of (a) mixing one or more candidate molecules with TGF-beta-binding-protein encoded by the nucleic acid molecule according to  
10 claim 1 and a selected member of the TGF-beta family of proteins (*e.g.*, BMP 5 or 6), (b) determining whether the candidate molecule alters the signaling of the TGF-beta family member, or alters the binding of the TGF-beta binding-protein to the TGF-beta family member. Within certain embodiments, the molecule alters the ability of TGF-beta to function as a positive regulator of mesenchymal cell differentiation. Within  
15 this aspect of the present invention, the candidate molecule(s) may alter signaling or binding by, for example, either decreasing (*e.g.*, inhibiting), or increasing (*e.g.*, enhancing) signaling or binding.

Within yet another aspect, methods are provided for determining whether a selected molecule is capable of increasing bone mineral content, comprising  
20 the step of determining whether a selected molecule inhibits the binding of TGF-beta binding-protein to bone, or an analogue thereof. Representative examples of bone or analogues thereof include hydroxyapatite and primary human bone samples obtained via biopsy.

Within certain embodiments of the above-recited methods, the selected  
25 molecule is contained within a mixture of molecules and the methods may further comprise the step of isolating one or more molecules which are functional within the assay. Within yet other embodiments, TGF-beta family of proteins is bound to a solid support and the binding of TGF-beta binding-protein is measured or TGF-beta binding-protein are bound to a solid support and the binding of TGF-beta proteins are measured.  
30

Utilizing methods such as those described above, a wide variety of molecules may be assayed for their ability to increase bone mineral content by inhibiting the binding of the TGF-beta binding-protein to the TGF-beta family of proteins. Representative examples of such molecules include proteins or peptides, organic molecules, and nucleic acid molecules.  
35

Within other related aspects of the invention, methods are provided for increasing bone mineral content in a warm-blooded animal, comprising the step of

administering to a warm-blooded animal a therapeutically effective amount of a molecule identified from the assays recited herein. Within another aspect, methods are provided for increasing bone mineral content in a warm-blooded animal, comprising the step of administering to a warm-blooded animal a therapeutically effective amount  
5 of a molecule which inhibits the binding of the TGF-beta binding-protein to the TGF-beta super-family of proteins, including bone morphogenic proteins (BMPs). Representative examples of suitable molecules include antisense molecules, ribozymes, ribozyme genes, and antibodies (e.g., a humanized antibody) which specifically recognize and alter the activity of the TGF-beta binding-protein.

10 Within another aspect of the present invention, methods are provided for increasing bone mineral content in a warm-blooded animal, comprising the steps of (a) introducing into cells which home to the bone a vector which directs the expression of a molecule which inhibits the binding of the TGF-beta binding-protein to the TGF-beta family of proteins and bone morphogenic proteins (BMPs), and (b) administering  
15 the vector-containing cells to a warm-blooded animal. As utilized herein, it should be understood that cells "home to bone" if they localize within the bone matrix after peripheral administration. Within one embodiment, such methods further comprise, prior to the step of introducing, isolating cells from the marrow of bone which home to the bone. Within a further embodiment, the cells which home to bone are selected from  
20 the group consisting of CD34+ cells and osteoblasts.

Within other aspects of the present invention, molecules are provided (preferably isolated) which inhibit the binding of the TGF-beta binding-protein to the TGF-beta super-family of proteins.

25 Within further embodiments, the molecules may be provided as a composition, and can further comprise an inhibitor of bone resorption. Representative examples of such inhibitors include calcitonin, estrogen, a bisphosphonate, a growth factor having anti-resorptive activity and tamoxifen.

30 Representative examples of molecules which may be utilized in the afore-mentioned therapeutic contexts include, e.g., ribozymes, ribozyme genes, antisense molecules, and/or antibodies (e.g., humanized antibodies). Such molecules may depending upon their selection, used to alter, antagonize, or agonize the signalling or binding of a TGF-beta binding-protein family member as described herein

35 Within various embodiments of the invention, the above-described molecules and methods of treatment or prevention may be utilized on conditions such as osteoporosis, osteomalasia, periodontal disease, scurvy, Cushing's Disease, bone fracture and conditions due to limb immobilization and steroid usage.

These and other aspects of the present invention will become evident upon reference to the following detailed description and attached drawings. In addition, various references are set forth herein which describe in more detail certain procedures or compositions (e.g., plasmids, etc.), and are therefore incorporated by  
5 reference in their entirety.

## BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 is a schematic illustration comparing the amino acid sequence of Human Dan; Human Gremlin; Human Cerberus and Human Beer. Arrows indicate the Cysteine backbone.

10 Figure 2 summarizes the results obtained from surveying a variety of human tissues for the expression of a TGF-beta binding-protein gene, specifically, the Human Beer gene. A semi-quantitative Reverse Transcription-Polymerase Chain Reaction (RT-PCR) procedure was used to amplify a portion of the gene from first-strand cDNA synthesized from total RNA (described in more detail in EXAMPLE 2A).

15 Figure 3 summarizes the results obtained from RNA *in situ* hybridization of mouse embryo sections, using a cRNA probe that is complementary to the mouse Beer transcript (described in more detail in EXAMPLE 2B). Panel A is a transverse section of 10.5 dpc embryo. Panel B is a sagittal section of 12.5 dpc embryo and panels C and D are sagittal sections of 15.5 dpc embryos.

20 Figure 4 illustrates, by western blot analysis, the specificity of three different polyclonal antibodies for their respective antigens (described in more detail in EXAMPLE 4). Figure 4A shows specific reactivity of an anti-H. Beer antibody for H. Beer antigen, but not H. Dan or H. Gremlin. Figure 4B shows reactivity of an anti-H. Gremlin antibody for H. Gremlin antigen, but not H. Beer or H. Dan. Figure 4C shows  
25 reactivity of an anti-H. Dan antibody for H. Dan, but not H. Beer or H. Gremlin.

Figure 5 illustrates, by western blot analysis, the selectivity of the TGF-beta binding-protein, Beer, for BMP-5 and BMP-6, but not BMP-4 (described in more detail in EXAMPLE 5).

30 Figure 6 demonstrates that the ionic interaction between the TGF-beta binding-protein, Beer, and BMP-5 has a dissociation constant in the 15-30 nM range.



## DETAILED DESCRIPTION OF THE INVENTION

## DEFINITIONS

Prior to setting forth the invention in detail, it may be helpful to an understanding thereof to set forth definitions of certain terms and to list and to define the abbreviations that will be used hereinafter.

"Molecule" should be understood to include proteins or peptides (*e.g.*, antibodies, recombinant binding partners, peptides with a desired binding affinity), nucleic acids (*e.g.*, DNA, RNA, chimeric nucleic acid molecules, and nucleic acid analogues such as PNA); and organic or inorganic compounds.

"TGF-beta" should be understood to include any known or novel member of the TGF-beta super-family, which also includes bone morphogenic proteins (BMPs).

"TGF-beta receptor" should be understood to refer to the receptor specific for a particular member of the TGF-beta super-family (including bone morphogenic proteins (BMPs)).

"TGF-beta binding-protein" should be understood to refer to a protein with specific binding affinity for a particular member or subset of members of the TGF-beta super-family (including bone morphogenic proteins (BMPs)). Specific examples of TGF-beta binding-proteins include proteins encoded by Sequence ID Nos. 1, 5, 7, 9, 11, 13, and 15.

Inhibiting the "binding of the TGF-beta binding-protein to the TGF-beta family of proteins and bone morphogenic proteins (BMPs)" should be understood to refer to molecules which allow the activation of TGF-beta or bone morphogenic proteins (BMPs), or allow the binding of TGF-beta family members including bone morphogenic proteins (BMPs) to their respective receptors, by removing or preventing TGF-beta from binding to TGF-binding-protein. Such inhibition may be accomplished, for example, by molecules which inhibit the binding of the TGF-beta binding-protein to specific members of the TGF-beta super-family.

"Vector" refers to an assembly which is capable of directing the expression of desired protein. The vector must include transcriptional promoter elements which are operably linked to the gene(s) of interest. The vector may be composed of either deoxyribonucleic acids ("DNA"), ribonucleic acids ("RNA"), or a combination of the two (*e.g.*, a DNA-RNA chimeric). Optionally, the vector may include a polyadenylation sequence, one or more restriction sites, as well as one or more selectable markers such as neomycin phosphotransferase or hygromycin phosphotransferase. Additionally, depending on the host cell chosen and the vector

employed, other genetic elements such as an origin of replication, additional nucleic acid restriction sites, enhancers, sequences conferring inducibility of transcription, and selectable markers, may also be incorporated into the vectors described herein.

An "isolated nucleic acid molecule" is a nucleic acid molecule that is not  
5 integrated in the genomic DNA of an organism. For example, a DNA molecule that encodes a TGF-binding protein that has been separated from the genomic DNA of a eukaryotic cell is an isolated DNA molecule. Another example of an isolated nucleic acid molecule is a chemically-synthesized nucleic acid molecule that is not integrated in the genome of an organism. The isolated nucleic acid molecule may be genomic DNA, cDNA,  
10 RNA, or composed at least in part of nucleic acid analogs.

An "isolated polypeptide" is a polypeptide that is essentially free from contaminating cellular components, such as carbohydrate, lipid, or other proteinaceous impurities associated with the polypeptide in nature. Within certain embodiments, a particular protein preparation contains an isolated polypeptide if it appears nominally as  
15 a single band on SDS-PAGE gel with Coomassie Blue staining. "Isolated" when referring to organic molecules means that the compounds are greater than 90 percent pure utilizing methods which are well known in the art (*e.g.*, NMR, melting point).

"Sclerosteosis" Sclerosteosis is a term that was applied by Hansen (1967) (Hansen, H. G., Sklerosteose. In: Opitz, H.; Schmid, F., Handbuch der  
20 Kinderheilkunde. Berlin: Springer (pub.) 6 1967. Pp. 351-355) to a disorder similar to van Buchem hyperostosis corticalis generalisata but possibly differing in radiologic appearance of the bone changes and in the presence of asymmetric cutaneous syndactyly of the index and middle fingers in many cases. The jaw has an unusually square appearance in this condition.

25 "Humanized antibodies" are recombinant proteins in which murine complementary determining regions of monoclonal antibodies have been transferred from heavy and light variable chains of the murine immunoglobulin into a human variable domain.

As used herein, an "antibody fragment" is a portion of an antibody such  
30 as F(ab')<sub>2</sub>, F(ab)<sub>2</sub>, Fab', Fab, and the like. Regardless of structure, an antibody fragment binds with the same antigen that is recognized by the intact antibody. For example, an anti-TGF-beta binding-protein monoclonal antibody fragment binds with an epitope of TGF-beta binding-protein.

The term "antibody fragment" also includes any synthetic or genetically  
35 engineered protein that acts like an antibody by binding to a specific antigen to form a complex. For example, antibody fragments include isolated fragments consisting of the

light chain variable region, "Fv" fragments consisting of the variable regions of the heavy and light chains, recombinant single chain polypeptide molecules in which light and heavy variable regions are connected by a peptide linker ("sFv proteins"), and minimal recognition units consisting of the amino acid residues that mimic the hypervariable region.

A "detectable label" is a molecule or atom which can be conjugated to an antibody moiety to produce a molecule useful for diagnosis. Examples of detectable labels include chelators, photoactive agents, radioisotopes, fluorescent agents, paramagnetic ions, enzymes, and other marker moieties.

As used herein, an "immunoconjugate" is a molecule comprising an anti-TGF-beta binding-protein antibody, or an antibody fragment, and a detectable label. An immunoconjugate has roughly the same, or only slightly reduced, ability to bind TGF-beta binding-protein after conjugation as before conjugation.

Abbreviations: TGF-beta - "Transforming Growth Factor-beta"; TGF-bBP - "Transforming Growth Factor-beta binding-protein" (one representative TGF-bBP is designated "H. Beer"); BMP - "bone morphogenic protein"; PCR - "polymerase chain reaction"; RT-PCR - PCR process in which RNA is first transcribed into DNA at the first step using reverse transcriptase (RT); cDNA - any DNA made by copying an RNA sequence into DNA form.

As noted above, the present invention provides a novel class of TGF-beta binding-proteins, as well as methods and compositions for increasing bone mineral content in warm-blooded animals. Briefly, the present inventions are based upon the unexpected discovery that a mutation in the gene which encodes a novel member of the TGF-beta binding-protein family results in a rare condition (sclerosteosis) characterized by bone mineral contents which are one- to four-fold higher than in normal individuals. Thus, as discussed in more detail below this discovery has led to the development of assays which may be utilized to select molecules which inhibit the binding of the TGF-beta binding-protein to the TGF-beta family of proteins and bone morphogenic proteins (BMPs), and methods of utilizing such molecules for increasing the bone mineral content of warm-blooded animals (including for example, humans).

#### DISCUSSION OF THE DISEASE KNOWN AS SCLEROSTEOSIS

Sclerosteosis is a term that was applied by Hansen (1967) (Hansen, H. G., Sklerosteose. In: Opitz, H.; Schmid, F., Handbuch der Kinderheilkunde. Berlin: Springer (pub.) 6 1967. Pp. 351-355) to a disorder similar to van Buchem hyperostosis

corticalis generalisata but possibly differing in radiologic appearance of the bone changes and in the presence of asymmetric cutaneous syndactyly of the index and middle fingers in many cases.

Sclerosteosis is now known to be an autosomal semi-dominant disorder which is characterized by widely disseminated sclerotic lesions of the bone in the adult. The condition is progressive. Sclerosteosis also has a developmental aspect which is associated with syndactyly (two or more fingers are fused together). The Sclerosteosis Syndrome is associated with large stature and many affected individuals attain a height of six feet or more. The bone mineral content of homozygotes can be 1 to 6 fold over normal individuals and bone mineral density can be 1 to 4 fold above normal values (e.g., from unaffected siblings).

The Sclerosteosis Syndrome occurs primarily in Afrikaans of Dutch descent in South Africa. Approximately 1/140 individuals in the Afrikaaner population are carriers of the mutated gene (heterozygotes). The mutation shows 100% penetrance. There are anecdotal reports of increased bone mineral density in heterozygotes with no associated pathologies (syndactyly or skull overgrowth).

It appears at the present time that there is no abnormality of the pituitary-hypothalamus axis in Sclerosteosis. In particular, there appears to be no overproduction of growth hormone and cortisone. In addition, sex hormone levels are normal in affected individuals. However, bone turnover markers (osteoblast specific alkaline phosphatase, osteocalcin, type 1 procollagen C' propeptide (PICP), and total alkaline phosphatase; (see Comier, C., *Curr. Opin. in Rheu.* 7:243, 1995) indicate that there is hyperosteoblastic activity associated with the disease but that there is normal to slightly decreased osteoclast activity as measured by markers of bone resorption (pyridinoline, deoxypyridinoline, N-telopeptide, urinary hydroxyproline, plasma tartrate-resistant acid phosphatases and galactosyl hydroxylysine (see Comier, *supra*)).

Sclerosteosis is characterized by the continual deposition of bone throughout the skeleton during the lifetime of the affected individuals. In homozygotes the continual deposition of bone mineral leads to an overgrowth of bone in areas of the skeleton where there is an absence of mechanoreceptors (skull, jaw, cranium). In homozygotes with Sclerosteosis, the overgrowth of the bones of the skull leads to cranial compression and eventually to death due to excessive hydrostatic pressure on the brain stem. In all other parts of the skeleton there is a generalized and diffuse sclerosis. Cortical areas of the long bones are greatly thickened resulting in a substantial increase in bone strength. Trabecular connections are increased in thickness

which in turn increases the strength of the trabecular bone. Sclerotic bones appear unusually opaque to x-rays.

As described in more detail in Example 1, the rare genetic mutation that is responsible for the Sclerosteosis syndrome has been localized to the region of human chromosome 17 that encodes a novel member of the TGF-beta binding-protein family (one representative example of which is designated "H. Beer"). As described in more detail below, based upon this discovery, the mechanism of bone mineralization is more fully understood, allowing the development of assays for molecules which increase bone mineralization, and use of such molecules to increase bone mineral content, and in the treatment or prevention of a wide number of diseases.

#### TGF-BETA SUPER-FAMILY

The Transforming Growth Factor-beta (TGF-beta) super-family contains a variety of growth factors that share common sequence elements and structural motifs (at both the secondary and tertiary levels). This protein family is known to exert a wide spectrum of biological responses on a large variety of cell types. Many of them have important functions during the embryonal development in pattern formation and tissue specification; in adults they are involved, *e.g.*, in wound healing and bone repair and bone remodeling, and in the modulation of the immune system. In addition to the three TGF-beta's, the super-family includes the Bone Morphogenic Proteins (BMPs), Activins, Inhibins, Growth and Differentiation Factors (GDFs), and Glial-Derived Neurotrophic Factors (GDNFs). Primary classification is established through general sequence features that bin a specific protein into a general sub-family. Additional stratification within the sub-family is possible due to stricter sequence conservation between members of the smaller group. In certain instances, such as with BMP-5, BMP-6 and BMP-7, this can be as high as 75 percent amino acid homology between members of the smaller group. This level of identity enables a single representative sequence to illustrate the key biochemical elements of the sub-group that separates it from other members of the larger family.

TGF-beta signals by inducing the formation of hetero-oligomeric complexes of type I and type II receptors. The crystal structure of TGF-beta2 has been determined. The general fold of the TGF-beta2 monomer contains a stable, compact, cysteine knotlike structure formed by three disulphide bridges. Dimerization, stabilized by one disulphide bridge, is antiparallel.

TGF-beta family members initiate their cellular action by binding to receptors with intrinsic serine/threonine kinase activity. This receptor family consists

of two subfamilies, denoted type I and type II receptors. Each member of the TGF-beta family binds to a characteristic combination of type I and type II receptors, both of which are needed for signaling. In the current model for TGF-beta receptor activation, TGF-beta first binds to the type II receptor (TbR-II), which occurs in the cell  
5 membrane in an oligomeric form with activated kinase. Thereafter, the type I receptor (TbR-I), which can not bind ligand in the absence of TbR-II, is recruited into the complex. TbR-II then phosphorylates TbR-I predominantly in a domain rich in glycine and serine residues (GS domain) in the juxtamembrane region, and thereby activates TbR-I.

10 Thus far seven type I receptors and five type II receptors have been identified.

BONE MORPHOGENIC PROTEINS (BMPs) ARE KEY REGULATORY PROTEINS IN  
DETERMINING BONE MINERAL DENSITY IN HUMANS

A major advance in the understanding of bone formation was the  
15 identification of the bone morphogenic proteins (BMPs), also known as osteogenic proteins (OPs), which regulate cartilage and bone differentiation in vivo. BMPs/OPs induce endochondral bone differentiation through a cascade of events which include formation of cartilage, hypertrophy and calcification of the cartilage, vascular invasion, differentiation of osteoblasts, and formation of bone. As described above, the  
20 BMPs/OPs (BMP 2-14, and osteogenic protein 1 and -2, OP-1 and OP-2) are members of the TGF-beta super-family. The striking evolutionary conservation between members the BMP/OP sub-family suggests that they are critical in the normal development and function of animals. Moreover, the presence of multiple forms of BMPs/OPs raises an important question about the biological relevance of this apparent  
25 redundancy. In addition to postfetal chondrogenesis and osteogenesis, the BMPs/OPs play multiple roles in skeletogenesis (including the development of craniofacial and dental tissues) and in embryonic development and organogenesis of parenchymatous organs, including the kidney. It is now understood that nature relies on common (and few) molecular mechanisms tailored to provide the emergence of specialized tissues  
30 and organs. The BMP/OP super-family is an elegant example of nature parsimony in programming multiple specialized functions deploying molecular isoforms with minor variation in amino acid motifs within highly conserved carboxy-terminal regions.

BMP ANTAGONISM

The BMP and Activin sub-families are subject to significant post-

translational regulation. An intricate extracellular control system exists, whereby a high affinity antagonist is synthesized and exported, and subsequently complexes selectively with BMPs or activins to disrupt their biological activity (W.C. Smith (1999) *TIG* 15(1) 3-6). A number of these natural antagonists have been identified, and  
5 based on sequence divergence appear to have evolved independently due to the lack of primary sequence conservation. There has been no structural work to date on this class of proteins. Studies of these antagonists has highlighted a distinct preference for interacting and neutralizing BMP-2 and BMP-4. Furthermore, the mechanism of inhibition seems to differ for the different antagonists (S. Iemura et al. (1998) *Proc*  
10 *Natl Acad Sci USA* 95 9337-9342).

#### NOVEL TGF-BETA BINDING-PROTEINS

##### 1. Background re: TGF-beta binding-proteins

As noted above, the present invention provides a novel class of TGF-beta binding-proteins that possess a nearly identical cysteine (disulfide) scaffold when  
15 compared to Human DAN, Human Gremlin, and Human Cerberus, and SCGF (U.S. Patent No. 5,780,263) but almost no homology at the nucleotide level (for background information, see generally Hsu, D.R., Economides, A.N., Wang, X., Eimon, P.M., Harland, R.M., "The *Xenopus* Dorsalizing Factor Gremlin Identifies a Novel Family of Secreted Proteins that Antagonize BMP Activities," *Molecular Cell* 1:673-683, 1998).

20 One representative example of the novel class of TGF-beta binding-proteins is disclosed in Sequence ID Nos. 1, 5, 9, 11, 13, and 15. Representative members of this class of binding proteins should also be understood to include variants of the TGF-beta binding-protein (e.g., Sequence ID Nos. 5 and 7). As utilized herein, a "TGF-beta binding-protein variant gene" refers to nucleic acid molecules that encode a  
25 polypeptide having an amino acid sequence that is a modification of SEQ ID Nos: 2, 10, 12, 14 or 16. Such variants include naturally-occurring polymorphisms or allelic variants of TGF-beta binding-protein genes, as well as synthetic genes that contain conservative amino acid substitutions of these amino acid sequences. Additional variant forms of a TGF-beta binding-protein gene are nucleic acid molecules that  
30 contain insertions or deletions of the nucleotide sequences described herein. TGF-beta binding-protein variant genes can be identified by determining whether the genes hybridize with a nucleic acid molecule having the nucleotide sequence of SEQ ID Nos: 1, 5, 7, 9, 11, 13, or 15 under stringent conditions. In addition, TGF-beta binding-protein variant genes should encode a protein having a cysteine backbone.

As an alternative, TGF-beta binding-protein variant genes can be identified by sequence comparison. As used herein, two amino acid sequences have "100% amino acid sequence identity" if the amino acid residues of the two amino acid sequences are the same when aligned for maximal correspondence. Similarly, two  
5 nucleotide sequences have "100% nucleotide sequence identity" if the nucleotide residues of the two nucleotide sequences are the same when aligned for maximal correspondence. Sequence comparisons can be performed using standard software programs such as those included in the LASERGENE bioinformatics computing suite, which is produced by DNASTAR (Madison, Wisconsin). Other methods for  
10 comparing two nucleotide or amino acid sequences by determining optimal alignment are well-known to those of skill in the art (see, for example, Peruski and Peruski, *The Internet and the New Biology: Tools for Genomic and Molecular Research* (ASM Press, Inc. 1997), Wu et al. (eds.), "Information Superhighway and Computer Databases of Nucleic Acids and Proteins," in *Methods in Gene Biotechnology*, pages  
15 123-151 (CRC Press, Inc. 1997), and Bishop (ed.), *Guide to Human Genome Computing*, 2nd Edition (Academic Press, Inc. 1998)).

A variant TGF-beta binding-protein should have at least a 50% amino acid sequence identity to SEQ ID NOs: 2, 6, 10, 12, 14 or 16 and preferably, greater than 60%, 65%, 70%, 75%, 80%, 85%, 90%, or 95% identity. Alternatively, TGF-beta  
20 binding-protein variants can be identified by having at least a 70% nucleotide sequence identity to SEQ ID NOs: 1, 5, 9, 11, 13 or 15. Moreover, the present invention contemplates TGF-beta binding-protein gene variants having greater than 75%, 80%, 85%, 90%, or 95% identity to SEQ ID NO:1. Regardless of the particular method used to identify a TGF-beta binding-protein variant gene or variant TGF-beta binding-  
25 protein, a variant TGF-beta binding-protein or a polypeptide encoded by a variant TGF-beta binding-protein gene can be functionally characterized by, for example, its ability to bind to and/or inhibit the signaling of a selected member of the TGF-beta family of proteins, or by its ability to bind specifically to an anti-TGF-beta binding-protein antibody.

30 The present invention includes functional fragments of TGF-beta binding-protein genes. Within the context of this invention, a "functional fragment" of a TGF-beta binding-protein gene refers to a nucleic acid molecule that encodes a portion of a TGF-beta binding-protein polypeptide which either (1) possesses the above-noted function activity, or (2) specifically binds with an anti-TGF-beta binding-  
35 protein antibody. For example, a functional fragment of a TGF-beta binding-protein gene described herein comprises a portion of the nucleotide sequence of SEQ ID Nos:



1, 5, 9, 11, 13, or 15.

## 2. Isolation of the TGF-beta binding-protein gene

DNA molecules encoding a binding-protein gene can be obtained by screening a human cDNA or genomic library using polynucleotide probes based upon, for example, SEQ ID NO: 1.

For example, the first step in the preparation of a cDNA library is to isolate RNA using methods well-known to those of skill in the art. In general, RNA isolation techniques must provide a method for breaking cells, a means of inhibiting RNase-directed degradation of RNA, and a method of separating RNA from DNA, protein, and polysaccharide contaminants. For example, total RNA can be isolated by freezing tissue in liquid nitrogen, grinding the frozen tissue with a mortar and pestle to lyse the cells, extracting the ground tissue with a solution of phenol/chloroform to remove proteins, and separating RNA from the remaining impurities by selective precipitation with lithium chloride (see, for example, Ausubel et al. (eds.), *Short Protocols in Molecular Biology*, 3rd Edition, pages 4-1 to 4-6 (John Wiley & Sons 1995) ["Ausubel (1995)"]; Wu et al., *Methods in Gene Biotechnology*, pages 33-41 (CRC Press, Inc. 1997) ["Wu (1997)"]).

Alternatively, total RNA can be isolated by extracting ground tissue with guanidinium isothiocyanate, extracting with organic solvents, and separating RNA from contaminants using differential centrifugation (see, for example, Ausubel (1995) at pages 4-1 to 4-6; Wu (1997) at pages 33-41).

In order to construct a cDNA library, poly(A)<sup>+</sup> RNA must be isolated from a total RNA preparation. Poly(A)<sup>+</sup> RNA can be isolated from total RNA by using the standard technique of oligo(dT)-cellulose chromatography (see, for example, Ausubel (1995) at pages 4-11 to 4-12).

Double-stranded cDNA molecules are synthesized from poly(A)<sup>+</sup> RNA using techniques well-known to those in the art. (see, for example, Wu (1997) at pages 41-46). Moreover, commercially available kits can be used to synthesize double-stranded cDNA molecules. For example, such kits are available from Life Technologies, Inc. (Gaithersburg, Maryland), CLONTECH Laboratories, Inc. (Palo Alto, California), Promega Corporation (Madison, Wisconsin) and Stratagene Cloning Systems (La Jolla, California).

The basic approach for obtaining TGF-beta binding-protein cDNA clones can be modified by constructing a subtracted cDNA library which is enriched in TGF-binding-protein-specific cDNA molecules. Techniques for constructing subtracted libraries are well-known to those of skill in the art (see, for example, Sargent, "Isolation of

Differentially Expressed Genes," in *Meth. Enzymol.* 152:423, 1987, and Wu et al. (eds.), "Construction and Screening of Subtracted and Complete Expression cDNA Libraries," in *Methods in Gene Biotechnology*, pages 29-65 (CRC Press, Inc. 1997)).

Various cloning vectors are appropriate for the construction of a cDNA library. For example, a cDNA library can be prepared in a vector derived from bacteriophage, such as a  $\lambda$ gt10 vector (see, for example, Huynh et al., "Constructing and Screening cDNA Libraries in  $\lambda$ gt10 and  $\lambda$ gt11," in *DNA Cloning: A Practical Approach Vol. I*, Glover (ed.), page 49 (IRL Press, 1985); Wu (1997) at pages 47-52).

Alternatively, double-stranded cDNA molecules can be inserted into a plasmid vector, such as a pBluescript vector (Stratagene Cloning Systems; La Jolla, California), a LambdaGEM-4 (Promega Corp.; Madison, Wisconsin) or other commercially available vectors. Suitable cloning vectors also can be obtained from the American Type Culture Collection (Rockville, Maryland).

In order to amplify the cloned cDNA molecules, the cDNA library is inserted into a prokaryotic host, using standard techniques. For example, a cDNA library can be introduced into competent *E. coli* DH5 cells, which can be obtained from Life Technologies, Inc. (Gaithersburg, Maryland).

A human genomic DNA library can be prepared by means well-known in the art (see, for example, Ausubel (1995) at pages 5-1 to 5-6; Wu (1997) at pages 307-327). Genomic DNA can be isolated by lysing tissue with the detergent Sarkosyl, digesting the lysate with proteinase K, clearing insoluble debris from the lysate by centrifugation, precipitating nucleic acid from the lysate using isopropanol, and purifying resuspended DNA on a cesium chloride density gradient.

DNA fragments that are suitable for the production of a genomic library can be obtained by the random shearing of genomic DNA or by the partial digestion of genomic DNA with restriction endonucleases. Genomic DNA fragments can be inserted into a vector, such as a bacteriophage or cosmid vector, in accordance with conventional techniques, such as the use of restriction enzyme digestion to provide appropriate termini, the use of alkaline phosphatase treatment to avoid undesirable joining of DNA molecules, and ligation with appropriate ligases. Techniques for such manipulation are well-known in the art (see, for example, Ausubel (1995) at pages 5-1 to 5-6; Wu (1997) at pages 307-327).

Nucleic acid molecules that encode a TGF-beta binding-protein gene can also be obtained using the polymerase chain reaction (PCR) with oligonucleotide primers having nucleotide sequences that are based upon the nucleotide sequences of the human TGF-beta binding-protein gene, as described herein. General methods for screening libraries with PCR are provided by, for example, Yu et al., "Use of the

Polymerase Chain Reaction to Screen Phage Libraries," in *Methods in Molecular Biology, Vol. 15: PCR Protocols: Current Methods and Applications*, White (ed.), pages 211-215 (Humana Press, Inc. 1993). Moreover, techniques for using PCR to isolate related genes are described by, for example, Preston, "Use of Degenerate  
5 Oligonucleotide Primers and the Polymerase Chain Reaction to Clone Gene Family Members," in *Methods in Molecular Biology, Vol. 15: PCR Protocols: Current Methods and Applications*, White (ed.), pages 317-337 (Humana Press, Inc. 1993).

Alternatively, human genomic libraries can be obtained from commercial sources such as Research Genetics (Huntsville, AL) and the American Type Culture  
10 Collection (Rockville, Maryland).

A library containing cDNA or genomic clones can be screened with one or more polynucleotide probes based upon SEQ ID NO:1, using standard methods (*see, for example*, Ausubel (1995) at pages 6-1 to 6-11).

Anti-TGF-beta binding-protein antibodies, produced as described below,  
15 can also be used to isolate DNA sequences that encode TGF-beta binding-protein genes from cDNA libraries. For example, the antibodies can be used to screen  $\lambda$ gt11 expression libraries, or the antibodies can be used for immunoscreening following hybrid selection and translation (*see, for example*, Ausubel (1995) at pages 6-12 to 6-16; Margolis et al., "Screening  $\lambda$  expression libraries with antibody and protein  
20 probes," in *DNA Cloning 2: Expression Systems, 2nd Edition*, Glover et al. (eds.), pages 1-14 (Oxford University Press 1995)).

The sequence of a TGF-beta binding-protein cDNA or TGF-beta binding-protein genomic fragment can be determined using standard methods. Moreover, the identification of genomic fragments containing a TGF-beta binding-  
25 protein promoter or regulatory element can be achieved using well-established techniques, such as deletion analysis (*see, generally*, Ausubel (1995)).

As an alternative, a TGF-beta binding-protein gene can be obtained by synthesizing DNA molecules using mutually priming long oligonucleotides and the nucleotide sequences described herein (*see, for example*, Ausubel (1995) at pages 8-8  
30 to 8-9). Established techniques using the polymerase chain reaction provide the ability to synthesize DNA molecules at least two kilobases in length (Adang et al., *Plant Molec. Biol.* 21:1131, 1993; Bambot et al., *PCR Methods and Applications* 2:266, 1993; Dillon et al., "Use of the Polymerase Chain Reaction for the Rapid Construction of Synthetic Genes," in *Methods in Molecular Biology, Vol. 15: PCR Protocols: Current Methods and Applications*, White (ed.), pages 263-268, (Humana Press, Inc.  
35 1993); Holowachuk et al., *PCR Methods Appl.* 4:299, 1995).

### 3. Production of TGF-beta binding-protein genes

Nucleic acid molecules encoding variant TGF-beta binding-protein genes can be obtained by screening various cDNA or genomic libraries with polynucleotide probes having nucleotide sequences based upon SEQ ID NO:1, 5, 9, 11, 13, or 15, using procedures described above. TGF-beta binding-protein gene variants can also be constructed synthetically. For example, a nucleic acid molecule can be devised that encodes a polypeptide having a conservative amino acid change, compared with the amino acid sequence of SEQ ID NOs: 2, 6, 8, 10, 12, 14, or 16. That is, variants can be obtained that contain one or more amino acid substitutions of SEQ ID NOs: 2, 6, 8, 10, 12, 14 or 16, in which an alkyl amino acid is substituted for an alkyl amino acid in a TGF-beta binding-protein amino acid sequence, an aromatic amino acid is substituted for an aromatic amino acid in a TGF-beta binding-protein amino acid sequence, a sulfur-containing amino acid is substituted for a sulfur-containing amino acid in a TGF-beta binding-protein amino acid sequence, a hydroxy-containing amino acid is substituted for a hydroxy-containing amino acid in a TGF-beta binding-protein amino acid sequence, an acidic amino acid is substituted for an acidic amino acid in a TGF-beta binding-protein amino acid sequence, a basic amino acid is substituted for a basic amino acid in a TGF-beta binding-protein amino acid sequence, or a dibasic monocarboxylic amino acid is substituted for a dibasic monocarboxylic amino acid in a TGF-beta binding-protein amino acid sequence.

Among the common amino acids, for example, a "conservative amino acid substitution" is illustrated by a substitution among amino acids within each of the following groups: (1) glycine, alanine, valine, leucine, and isoleucine, (2) phenylalanine, tyrosine, and tryptophan, (3) serine and threonine, (4) aspartate and glutamate, (5) glutamine and asparagine, and (6) lysine, arginine and histidine. In making such substitutions, it is important to, where possible, maintain the cysteine backbone outlined in Figure 1.

Conservative amino acid changes in a TGF-beta binding-protein gene can be introduced by substituting nucleotides for the nucleotides recited in SEQ ID NO:1. Such "conservative amino acid" variants can be obtained, for example, by oligonucleotide-directed mutagenesis, linker-scanning mutagenesis, mutagenesis using the polymerase chain reaction, and the like (see Ausubel (1995) at pages 8-10 to 8-22; and McPherson (ed.), *Directed Mutagenesis: A Practical Approach* (IRL Press 1991)). The functional ability of such variants can be determined using a standard method, such as the assay described herein. Alternatively, a variant TGF-beta binding-protein polypeptide can be identified by the ability to specifically bind anti-TGF-beta binding-

protein antibodies.

Routine deletion analyses of nucleic acid molecules can be performed to obtain "functional fragments" of a nucleic acid molecule that encodes a TGF-beta binding-protein polypeptide. As an illustration, DNA molecules having the nucleotide sequence of SEQ ID NO:1 can be digested with *Bal31* nuclease to obtain a series of nested deletions. The fragments are then inserted into expression vectors in proper reading frame, and the expressed polypeptides are isolated and tested for activity, or for the ability to bind anti-TGF-beta binding-protein antibodies. One alternative to exonuclease digestion is to use oligonucleotide-directed mutagenesis to introduce deletions or stop codons to specify production of a desired fragment. Alternatively, particular fragments of a TGF-beta binding-protein gene can be synthesized using the polymerase chain reaction.

Standard techniques for functional analysis of proteins are described by, for example, Treuter et al., *Molec. Gen. Genet.* 240:113, 1993; Content et al., "Expression and preliminary deletion analysis of the 42 kDa 2-5A synthetase induced by human interferon," in *Biological Interferon Systems, Proceedings of ISIR-TNO Meeting on Interferon Systems*, Cantell (ed.), pages 65-72 (Nijhoff 1987); Herschman, "The EGF Receptor," in *Control of Animal Cell Proliferation, Vol. 1*, Boynton et al., (eds.) pages 169-199 (Academic Press 1985); Coumailleau et al., *J. Biol. Chem.* 270:29270, 1995; Fukunaga et al., *J. Biol. Chem.* 270:25291, 1995; Yamaguchi et al., *Biochem. Pharmacol.* 50:1295, 1995; and Meisel et al., *Plant Molec. Biol.* 30:1, 1996.

The present invention also contemplates functional fragments of a TGF-beta binding-protein gene that have conservative amino acid changes.

A TGF-beta binding-protein variant gene can be identified on the basis of structure by determining the level of identity with nucleotide and amino acid sequences of SEQ ID NOs: 1, 5, 9, 11, 13, or, 15 and 2, 6, 10, 12, 14, or 16, as discussed above. An alternative approach to identifying a variant gene on the basis of structure is to determine whether a nucleic acid molecule encoding a potential variant TGF-beta binding-protein gene can hybridize under stringent conditions to a nucleic acid molecule having the nucleotide sequence of SEQ ID Nos: 1, 5, 9, 11, 13, or, 15, or a portion thereof of at least 15 or 20 nucleotides in length. As an illustration of stringent hybridization conditions, a nucleic acid molecule having a variant TGF-beta binding-protein sequence can bind with a fragment of a nucleic acid molecule having a sequence from SEQ ID NO:1 in a buffer containing, for example, 5xSSPE (1xSSPE = 180 mM sodium chloride, 10 mM sodium phosphate, 1 mM EDTA (pH 7.7), 5xDenhardt's solution (100xDenhardt's = 2% (w/v) bovine serum albumin, 2% (w/v)

Ficoll, 2% (w/v) polyvinylpyrrolidone) and 0.5% SDS incubated overnight at 55-60°C. Post-hybridization washes at high stringency are typically performed in 0.5xSSC (1xSSC = 150 mM sodium chloride, 15 mM trisodium citrate) or in 0.5xSSPE at 55-60°C.

5           Regardless of the particular nucleotide sequence of a variant TGF-beta binding-protein gene, the gene encodes a polypeptide that can be characterized by its functional activity, or by the ability to bind specifically to an anti-TGF-beta binding-protein antibody. More specifically, variant TGF-beta binding-protein genes encode polypeptides which exhibit at least 50%, and preferably, greater than 60, 70, 80 or  
10 90%, of the activity of polypeptides encoded by the human TGF-beta binding-protein gene described herein.

#### 4.       Production of TGF-beta binding-protein in Cultured Cells

To express a TGF-beta binding-protein gene, a nucleic acid molecule encoding the polypeptide must be operably linked to regulatory sequences that control  
15 transcriptional expression in an expression vector and then introduced into a host cell. In addition to transcriptional regulatory sequences, such as promoters and enhancers, expression vectors can include translational regulatory sequences and a marker gene which is suitable for selection of cells that carry the expression vector.

Expression vectors that are suitable for production of a foreign protein in  
20 eukaryotic cells typically contain (1) prokaryotic DNA elements coding for a bacterial replication origin and an antibiotic resistance marker to provide for the growth and selection of the expression vector in a bacterial host; (2) eukaryotic DNA elements that control initiation of transcription, such as a promoter; and (3) DNA elements that control the processing of transcripts, such as a transcription termination/polyadenylation sequence.

25           TGF-beta binding-proteins of the present invention are preferably expressed in mammalian cells. Examples of mammalian host cells include African green monkey kidney cells (Vero; ATCC CRL 1587), human embryonic kidney cells (293-HEK; ATCC CRL 1573), baby hamster kidney cells (BHK-21; ATCC CRL 8544), canine kidney cells (MDCK; ATCC CCL 34), Chinese hamster ovary cells (CHO-K1; ATCC CCL61), rat  
30 pituitary cells (GH1; ATCC CCL82), HeLa S3 cells (ATCC CCL2.2), rat hepatoma cells (H-4-II-E; ATCC CRL 1548) SV40-transformed monkey kidney cells (COS-1; ATCC CRL 1650) and murine embryonic cells (NIH-3T3; ATCC CRL 1658).

For a mammalian host, the transcriptional and translational regulatory signals may be derived from viral sources, such as adenovirus, bovine papilloma virus,  
35 simian virus, or the like, in which the regulatory signals are associated with a particular gene

which has a high level of expression. Suitable transcriptional and translational regulatory sequences also can be obtained from mammalian genes, such as actin, collagen, myosin, and metallothionein genes.

Transcriptional regulatory sequences include a promoter region sufficient to  
 5 direct the initiation of RNA synthesis. Suitable eukaryotic promoters include the promoter of the mouse metallothionein I gene [Hamer et al., *J. Molec. Appl. Genet.* 1:273, 1982], the TK promoter of *Herpes* virus [McKnight, *Cell* 31:355, 1982], the *SV40* early promoter [Benoist et al., *Nature* 290:304, 1981], the *Rous* sarcoma virus promoter [Gorman et al., *Proc. Nat'l Acad. Sci. USA* 79:6777, 1982], the cytomegalovirus promoter [Foecking et al.,  
 10 *Gene* 45:101, 1980], and the mouse mammary tumor virus promoter (see, generally, Etcheverry, "Expression of Engineered Proteins in Mammalian Cell Culture," in *Protein Engineering: Principles and Practice*, Cleland et al. (eds.), pages 163-181 (John Wiley & Sons, Inc. 1996)).

Alternatively, a prokaryotic promoter, such as the bacteriophage T3 RNA  
 15 polymerase promoter, can be used to control TGF-beta binding-protein gene expression in mammalian cells if the prokaryotic promoter is regulated by a eukaryotic promoter (Zhou et al., *Mol. Cell. Biol.* 10:4529, 1990; Kaufman et al., *Nucl. Acids Res.* 19:4485, 1991).

TGF-beta binding-protein genes may also be expressed in bacterial, yeast, insect, or plant cells. Suitable promoters that can be used to express TGF-beta binding-  
 20 protein polypeptides in a prokaryotic host are well-known to those of skill in the art and include promoters capable of recognizing the T4, T3, Sp6 and T7 polymerases, the  $P_R$  and  $P_L$  promoters of bacteriophage lambda, the *trp*, *recA*, heat shock, *lacUV5*, *tac*, *lpp-lacSpr*, *phoA*, and *lacZ* promoters of *E. coli*, promoters of *B. subtilis*, the promoters of the bacteriophages of *Bacillus*, *Streptomyces* promoters, the *int* promoter of bacterio-  
 25 phage lambda, the *bla* promoter of pBR322, and the CAT promoter of the chloramphenicol acetyl transferase gene. Prokaryotic promoters have been reviewed by Glick, *J. Ind. Microbiol.* 1:277, 1987, Watson et al., *Molecular Biology of the Gene*, 4th Ed. (Benjamin Cummins 1987), and by Ausubel et al. (1995).

Preferred prokaryotic hosts include *E. coli* and *Bacillus subtilis*.  
 30 Suitable strains of *E. coli* include BL21(DE3), BL21(DE3)pLysS, BL21(DE3)pLysE, DH1, DH4I, DH5, DH5I, DH5IF', DH5IMCR, DH10B, DH10B/p3, DH11S, C600, HB101, JM101, JM105, JM109, JM110, K38, RR1, Y1088, Y1089, CSH18, ER145I, and ER1647 (see, for example, Brown (Ed.), *Molecular Biology Labfax* (Academic Press 1991)). Suitable strains of *Bacillus subtilis* include BR151, YB886, M1119,  
 35 M1120, and B170 (see, for example, Hardy, "Bacillus Cloning Methods," in *DNA Cloning: A Practical Approach*, Glover (Ed.) (IRL Press 1985)).

Methods for expressing proteins in prokaryotic hosts are well-known to those of skill in the art (see, for example, Williams et al., "Expression of foreign proteins in *E. coli* using plasmid vectors and purification of specific polyclonal antibodies," in *DNA Cloning 2: Expression Systems, 2nd Edition*, Glover et al. (eds.), page 15 (Oxford University Press 1995); Ward et al., "Genetic Manipulation and Expression of Antibodies," in *Monoclonal Antibodies: Principles and Applications*, page 137 (Wiley-Liss, Inc. 1995); and Georgiou, "Expression of Proteins in Bacteria," in *Protein Engineering: Principles and Practice*, Cleland et al. (eds.), page 101 (John Wiley & Sons, Inc. 1996)).

10 The baculovirus system provides an efficient means to introduce cloned *TGF-beta binding-protein* genes into insect cells. Suitable expression vectors are based upon the *Autographa californica* multiple nuclear polyhedrosis virus (AcMNPV), and contain well-known promoters such as *Drosophila* heat shock protein (hsp) 70 promoter, *Autographa californica* nuclear polyhedrosis virus immediate-early gene  
15 promoter (*ie-1*) and the delayed early 39K promoter, baculovirus p10 promoter, and the *Drosophila* metallothionein promoter. Suitable insect host cells include cell lines derived from IPLB-Sf-21, a *Spodoptera frugiperda* pupal ovarian cell line, such as Sf9 (ATCC CRL 1711), Sf21AE, and Sf21 (Invitrogen Corporation, San Diego, CA), as well as *Drosophila* Schneider-2 cells. Established techniques for producing  
20 recombinant proteins in baculovirus systems are provided by Bailey et al., "Manipulation of Baculovirus Vectors," in *Methods in Molecular Biology, Volume 7: Gene Transfer and Expression Protocols*, Murray (ed.), pages 147-168 (The Humana Press, Inc. 1991), by Patel et al., "The baculovirus expression system," in *DNA Cloning 2: Expression Systems, 2nd Edition*, Glover et al. (eds.), pages 205-244 (Oxford University Press 1995), by Ausubel (1995) at pages 16-37 to 16-57, by Richardson (ed.), *Baculovirus Expression Protocols* (The Humana Press, Inc. 1995), and by Lucknow, "Insect Cell Expression Technology," in *Protein Engineering: Principles and Practice*, Cleland et al. (eds.), pages 183-218 (John Wiley & Sons, Inc. 1996).

Promoters for expression in yeast include promoters from *GAL1* (galactose), *PGK* (phosphoglycerate kinase), *ADH* (alcohol dehydrogenase), *AOX1* (alcohol oxidase), *HIS4* (histidinol dehydrogenase), and the like. Many yeast cloning vectors have been designed and are readily available. These vectors include YIp-based vectors, such as YIp5, YRp vectors, such as YRp17, YEpl3 and YCp vectors, such as YCp19. One skilled in the art will appreciate that there are a  
35 wide variety of suitable vectors for expression in yeast cells.

Expression vectors can also be introduced into plant protoplasts, intact plant



tissues, or isolated plant cells. General methods of culturing plant tissues are provided, for example, by Miki et al., "Procedures for Introducing Foreign DNA into Plants," in *Methods in Plant Molecular Biology and Biotechnology*, Glick et al. (eds.), pages 67-88 (CRC Press, 1993).

5 An expression vector can be introduced into host cells using a variety of standard techniques including calcium phosphate transfection, liposome-mediated transfection, microprojectile-mediated delivery, electroporation, and the like. Preferably, the transfected cells are selected and propagated to provide recombinant host cells that  
10 introduce vectors into eukaryotic cells and techniques for selecting such stable transformants using a dominant selectable marker are described, for example, by Ausubel (1995) and by Murray (ed.), *Gene Transfer and Expression Protocols* (Humana Press 1991). Methods for introducing expression vectors into bacterial, yeast, insect, and plant cells are also provided by Ausubel (1995).

15 General methods for expressing and recovering foreign protein produced by a mammalian cell system is provided by, for example, Etcheverry, "Expression of Engineered Proteins in Mammalian Cell Culture," in *Protein Engineering: Principles and Practice*, Cleland et al. (eds.), pages 163 (Wiley-Liss, Inc. 1996). Standard techniques for recovering protein produced by a bacterial system is provided by, for example,  
20 Grisshammer et al., "Purification of over-produced proteins from *E. coli* cells," in *DNA Cloning 2: Expression Systems, 2nd Edition*, Glover et al. (eds.), pages 59-92 (Oxford University Press 1995). Established methods for isolating recombinant proteins from a baculovirus system are described by Richardson (ed.), *Baculovirus Expression Protocols* (The Humana Press, Inc., 1995).

25 More generally, TGF-beta binding-protein can be isolated by standard techniques, such as affinity chromatography, size exclusion chromatography, ion exchange chromatography, HPLC and the like. Additional variations in TGF-beta binding-protein isolation and purification can be devised by those of skill in the art. For example, anti-TGF-beta binding-protein antibodies, obtained as described below,  
30 can be used to isolate large quantities of protein by immunoaffinity purification.

#### 5. Production of Antibodies to TGF-beta binding-proteins

Antibodies to TGF-beta binding-protein can be obtained, for example, using the product of an expression vector as an antigen. Particularly useful anti-TGF-beta binding-protein antibodies "bind specifically" with TGF-beta binding-protein of  
35 Sequence ID Nos. 2, 6, 10, 12, 14, or 16, but not to other TGF-beta binding-proteins

such as Dan, Cerberus, SCGF, or Gremlin. Antibodies of the present invention (including fragments and derivatives thereof) may be a polyclonal or, especially a monoclonal antibody. The antibody may belong to any immunoglobulin class, and may be for example an IgG, for example IgG<sub>1</sub>, IgG<sub>2</sub>, IgG<sub>3</sub>, IgG<sub>4</sub>, IgE, IgM, or IgA antibody. It  
5 may be of animal, for example mammalian origin, and may be for example a murine, rat, human or other primate antibody. Where desired the antibody may be an internalising antibody.

Polyclonal antibodies to recombinant TGF-beta binding-protein can be prepared using methods well-known to those of skill in the art (see, for example, Green  
10 et al., "Production of Polyclonal Antisera," in *Immunochemical Protocols* (Manson, ed.), pages 1-5 (Humana Press 1992); Williams et al., "Expression of foreign proteins in *E. coli* using plasmid vectors and purification of specific polyclonal antibodies," in *DNA Cloning 2: Expression Systems, 2nd Edition*, Glover et al. (eds.), page 15 (Oxford University Press 1995)). Although polyclonal antibodies are typically raised in animals  
15 such as rats, mice, rabbits, goats, or sheep, an anti-TGF-beta binding-protein antibody of the present invention may also be derived from a subhuman primate antibody. General techniques for raising diagnostically and therapeutically useful antibodies in baboons may be found, for example, in Goldenberg et al., international patent publication No. WO 91/11465 (1991), and in Losman et al., *Int. J. Cancer* 46:310,  
20 1990.

The antibody should comprise at least a variable region domain. The variable region domain may be of any size or amino acid composition and will generally comprise at least one hypervariable amino acid sequence responsible for antigen binding embedded in a framework sequence. In general terms the variable (V) region domain may  
25 be any suitable arrangement of immunoglobulin heavy (V<sub>H</sub>) and/or light (V<sub>L</sub>) chain variable domains. Thus for example the V region domain may be monomeric and be a V<sub>H</sub> or V<sub>L</sub> domain where these are capable of independently binding antigen with acceptable affinity. Alternatively the V region domain may be dimeric and contain V<sub>H</sub>-V<sub>H</sub>, V<sub>H</sub>-V<sub>L</sub>, or V<sub>L</sub>-V<sub>L</sub> dimers in which the V<sub>H</sub> and V<sub>L</sub> chains are non-covalently associated (abbreviated  
30 hereinafter as F<sub>v</sub>). Where desired, however, the chains may be covalently coupled either directly, for example via a disulphide bond between the two variable domains, or through a linker, for example a peptide linker, to form a single chain domain (abbreviated hereinafter as scF<sub>v</sub>).

The variable region domain may be any naturally occurring variable domain  
35 or an engineered version thereof. By engineered version is meant a variable region domain which has been created using recombinant DNA engineering techniques. Such engineered

versions include those created for example from natural antibody variable regions by insertions, deletions or changes in or to the amino acid sequences of the natural antibodies. Particular examples of this type include those engineered variable region domains containing at least one CDR and optionally one or more framework amino acids from one antibody and the remainder of the variable region domain from a second antibody.

The variable region domain may be covalently attached at a C-terminal amino acid to at least one other antibody domain or a fragment thereof. Thus, for example where a  $V_H$  domain is present in the variable region domain this may be linked to an immunoglobulin  $C_H1$  domain or a fragment thereof. Similarly a  $V_L$  domain may be linked to a  $C_K$  domain or a fragment thereof. In this way for example the antibody may be a Fab fragment wherein the antigen binding domain contains associated  $V_H$  and  $V_L$  domains covalently linked at their C-termini to a  $CH1$  and  $C_K$  domain respectively. The  $CH1$  domain may be extended with further amino acids, for example to provide a hinge region domain as found in a Fab' fragment, or to provide further domains, such as antibody  $CH2$  and  $CH3$  domains.

Another form of an antibody fragment is a peptide coding for a single complementarity-determining region (CDR). CDR peptides ("minimal recognition units") can be obtained by constructing genes encoding the CDR of an antibody of interest. Such genes are prepared, for example, by using the polymerase chain reaction to synthesize the variable region from RNA of antibody-producing cells (see, for example, Larrick et al., *Methods: A Companion to Methods in Enzymology* 2:106, 1991; Courtenay-Luck, "Genetic Manipulation of Monoclonal Antibodies," in *Monoclonal Antibodies: Production, Engineering and Clinical Application*, Ritter et al. (eds.), page 166 (Cambridge University Press 1995); and Ward et al., "Genetic Manipulation and Expression of Antibodies," in *Monoclonal Antibodies: Principles and Applications*, Birch et al., (eds.), page 137 (Wiley-Liss, Inc. 1995)).

Antibodies for use in the invention may in general be monoclonal (prepared by conventional immunisation and cell fusion procedures) or in the case of fragments, derived therefrom using any suitable standard chemical e.g. reduction or enzymatic cleavage and/or digestion techniques, for example by treatment with pepsin.

More specifically, monoclonal anti-TGF-beta binding-protein antibodies can be generated utilizing a variety of techniques. Rodent monoclonal antibodies to specific antigens may be obtained by methods known to those skilled in the art (see, for example, Kohler et al., *Nature* 256:495, 1975; and Coligan et al. (eds.), *Current Protocols in Immunology*, 1:2.5.1-2.6.7 (John Wiley & Sons 1991) ["Coligan"]; Picksley et al., "Production of monoclonal antibodies against proteins expressed in *E.*

*coli*," in *DNA Cloning 2: Expression Systems, 2nd Edition*, Glover et al. (eds.), page 93 (Oxford University Press 1995)).

Briefly, monoclonal antibodies can be obtained by injecting mice with a composition comprising a TGF-beta binding-protein gene product, verifying the presence of antibody production by removing a serum sample, removing the spleen to obtain B-lymphocytes, fusing the B-lymphocytes with myeloma cells to produce hybridomas, cloning the hybridomas, selecting positive clones which produce antibodies to the antigen, culturing the clones that produce antibodies to the antigen, and isolating the antibodies from the hybridoma cultures.

In addition, an anti-TGF-beta binding-protein antibody of the present invention may be derived from a human monoclonal antibody. Human monoclonal antibodies are obtained from transgenic mice that have been engineered to produce specific human antibodies in response to antigenic challenge. In this technique, elements of the human heavy and light chain locus are introduced into strains of mice derived from embryonic stem cell lines that contain targeted disruptions of the endogenous heavy chain and light chain loci. The transgenic mice can synthesize human antibodies specific for human antigens, and the mice can be used to produce human antibody-secreting hybridomas. Methods for obtaining human antibodies from transgenic mice are described, for example, by Green et al., *Nature Genet.* 7:13, 1994; Lonberg et al., *Nature* 368:856, 1994; and Taylor et al., *Int. Immun.* 6:579, 1994.

Monoclonal antibodies can be isolated and purified from hybridoma cultures by a variety of well-established techniques. Such isolation techniques include affinity chromatography with Protein-A Sepharose, size-exclusion chromatography, and ion-exchange chromatography (see, for example, Coligan at pages 2.7.1-2.7.12 and pages 2.9.1-2.9.3; Baines et al., "Purification of Immunoglobulin G (IgG)," in *Methods in Molecular Biology, Vol. 10*, pages 79-104 (The Humana Press, Inc. 1992)).

For particular uses, it may be desirable to prepare fragments of anti-TGF-beta binding-protein antibodies. Such antibody fragments can be obtained, for example, by proteolytic hydrolysis of the antibody. Antibody fragments can be obtained by pepsin or papain digestion of whole antibodies by conventional methods. As an illustration, antibody fragments can be produced by enzymatic cleavage of antibodies with pepsin to provide a 5S fragment denoted  $F(ab')_2$ . This fragment can be further cleaved using a thiol reducing agent to produce 3.5S Fab' monovalent fragments. Optionally, the cleavage reaction can be performed using a blocking group for the sulfhydryl groups that result from cleavage of disulfide linkages. As an alternative, an enzymatic cleavage using pepsin produces two monovalent Fab

fragments and an Fc fragment directly. These methods are described, for example, by Goldenberg, U.S. patent No. 4,331,647, Nisonoff et al., *Arch Biochem. Biophys.* 89:230, 1960, Porter, *Biochem. J.* 73:119, 1959, Edelman et al., in *Methods in Enzymology* 1:422 (Academic Press 1967), and by Coligan at pages 2.8.1-2.8.10 and 2.10.-2.10.4.

Other methods of cleaving antibodies, such as separation of heavy chains to form monovalent light-heavy chain fragments, further cleavage of fragments, or other enzymatic, chemical or genetic techniques may also be used, so long as the fragments bind to the antigen that is recognized by the intact antibody.

Alternatively, the antibody may be a recombinant or engineered antibody obtained by the use of recombinant DNA techniques involving the manipulation and re-expression of DNA encoding antibody variable and/or constant regions. Such DNA is known and/or is readily available from DNA libraries including for example phage-antibody libraries (see Chiswell, D J and McCafferty, J. *Tibtech.* 10 80-84 (1992)) or where desired can be synthesised. Standard molecular biology and/or chemistry procedures may be used to sequence and manipulate the DNA, for example, to introduce codons to create cysteine residues, to modify, add or delete other amino acids or domains as desired.

From here, one or more replicable expression vectors containing the DNA may be prepared and used to transform an appropriate cell line, e.g. a non-producing myeloma cell line, such as a mouse NSO line or a bacterial, e.g. *E.coli* line, in which production of the antibody will occur. In order to obtain efficient transcription and translation, the DNA sequence in each vector should include appropriate regulatory sequences, particularly a promoter and leader sequence operably linked to the variable domain sequence. Particular methods for producing antibodies in this way are generally well known and routinely used. For example, basic molecular biology procedures are described by Maniatis *et al* (Molecular Cloning, Cold Spring Harbor Laboratory, New York, 1989); DNA sequencing can be performed as described in Sanger *et al* (PNAS 74, 5463, (1977)) and the Amersham International plc sequencing handbook; and site directed mutagenesis can be carried out according to the method of Kramer *et al* (Nucl. Acids Res. 12, 9441, (1984)) and the Anglian Biotechnology Ltd handbook. Additionally, there are numerous publications, detailing techniques suitable for the preparation of antibodies by manipulation of DNA, creation of expression vectors and transformation of appropriate cells, for example as reviewed by Mountain A and Adair, J R in *Biotechnology and Genetic Engineering Reviews* (ed. Tombs, M P, 10, Chapter 1, 1992, Intercept, Andover, UK) and in International Patent Specification No. WO 91/09967.

Where desired, the antibody according to the invention may have one or

more effector or reporter molecules attached to it and the invention extends to such modified proteins. The effector or reporter molecules may be attached to the antibody through any available amino acid side-chain, terminal amino acid or, where present carbohydrate functional group located in the antibody, always provided of course that this  
5 does not adversely affect the binding properties and eventual usefulness of the molecule. Particular functional groups include, for example any free amino, imino, thiol, hydroxyl, carboxyl or aldehyde group. Attachment of the antibody and the effector and/or reporter molecule(s) may be achieved via such groups and an appropriate functional group in the effector or reporter molecules. The linkage may be direct or indirect, through spacing or  
10 bridging groups.

Effector molecules include, for example, antineoplastic agents, toxins (such as enzymatically active toxins of bacterial or plant origin and fragments thereof e.g. ricin and fragments thereof) biologically active proteins, for example enzymes, nucleic acids and fragments thereof, e.g. DNA, RNA and fragments thereof, naturally occurring and synthetic  
15 polymers e.g. polysaccharides and polyalkylene polymers such as poly(ethylene glycol) and derivatives thereof, radionuclides, particularly radioiodide, and chelated metals. Suitable reporter groups include chelated metals, fluorescent compounds or compounds which may be detected by NMR or ESR spectroscopy.

Particular antineoplastic agents include cytotoxic and cytostatic agents, for  
20 example alkylating agents, such as nitrogen mustards (e.g. chlorambucil, melphalan, mechlorethamine, cyclophosphamide, or uracil mustard) and derivatives thereof, triethylenephosphoramidate, triethylenethiophosphor-amide, busulphan, or cisplatin; antimetabolites, such as methotrexate, fluorouracil, floxuridine, cytarabine, mercaptopurine, thioguanine, fluoroacetic acid or fluorocitric acid, antibiotics, such as bleomycins (e.g.  
25 bleomycin sulphate), doxorubicin, daunorubicin, mitomycins (e.g. mitomycin C), actinomycins (e.g. dactinomycin) plicamycin, calichaemicin and derivatives thereof, or esperamicin and derivatives thereof; mitotic inhibitors, such as etoposide, vincristine or vinblastine and derivatives thereof; alkaloids, such as ellipticine; polyols such as taxicin-I or taxicin-II; hormones, such as androgens (e.g. dromostanolone or testolactone), progestins  
30 (e.g. megestrol acetate or medroxyprogesterone acetate), estrogens (e.g. dimethylstilbestrol diphosphate, polyestradiol phosphate or estramustine phosphate) or antiestrogens (e.g. tamoxifen); anthraquinones, such as mitoxantrone, ureas, such as hydroxyurea; hydrazines, such as procarbazine; or imidazoles, such as dacarbazine.

Particularly useful effector groups are calichaemicin and derivatives thereof  
35 (see for example South African Patent Specifications Nos. 85/8794, 88/8127 and 90/2839).

Chelated metals include chelates of di-or tripositive metals having a

coordination number from 2 to 8 inclusive. Particular examples of such metals include technetium (Tc), rhenium (Re), cobalt (Co), copper (Cu), gold (Au), silver (Ag), lead (Pb), bismuth (Bi), indium (In), gallium (Ga), yttrium (Y), terbium (Tb), gadolinium (Gd), and scandium (Sc). In general the metal is preferably a radionuclide. Particular radionuclides include <sup>99m</sup>Tc, <sup>186</sup>Re, <sup>188</sup>Re, <sup>58</sup>Co, <sup>60</sup>Co, <sup>67</sup>Cu, <sup>195</sup>Au, <sup>199</sup>Au, <sup>110</sup>Ag, <sup>203</sup>Pb, <sup>206</sup>Bi, <sup>207</sup>Bi, <sup>111</sup>In, <sup>67</sup>Ga, <sup>68</sup>Ga, <sup>88</sup>Y, <sup>90</sup>Y, <sup>160</sup>Tb, <sup>153</sup>Gd and <sup>47</sup>Sc.

The chelated metal may be for example one of the above types of metal chelated with any suitable polydentate chelating agent, for example acyclic or cyclic polyamines, polyethers, (e.g. crown ethers and derivatives thereof); polyamides; porphyrins; and carbocyclic derivatives.

In general, the type of chelating agent will depend on the metal in use. One particularly useful group of chelating agents in conjugates according to the invention, however, are acyclic and cyclic polyamines, especially polyaminocarboxylic acids, for example diethylenetriaminepentaacetic acid and derivatives thereof, and macrocyclic amines, e.g. cyclic tri-aza and tetra-aza derivatives (for example as described in International Patent Specification No. WO 92/22583); and polyamides, especially desferrioxamine and derivatives thereof.

Thus for example when it is desired to use a thiol group in the antibody as the point of attachment this may be achieved through reaction with a thiol reactive group present in the effector or reporter molecule. Examples of such groups include an  $\alpha$ -halocarboxylic acid or ester, e.g. iodoacetamide, an imide, e.g. maleimide, a vinyl sulphone, or a disulphide. These and other suitable linking procedures are generally and more particularly described in International Patent Specifications Nos. WO 93/06231, WO 92/22583, WO 90/091195 and WO 89/01476.

#### 25      ASSAYS FOR SELECTING MOLECULES WHICH INCREASE BONE DENSITY

As discussed above, the present invention provides methods for selecting and/or isolating compounds which are capable of increasing bone density. For example, within one aspect of the present invention methods are provided for determining whether a selected molecule is capable of increasing bone mineral content, comprising the steps of (a) mixing a selected molecule with TGF-beta binding protein and a selected member of the TGF-beta family of proteins, (b) determining whether the selected molecule stimulates signaling by the TGF-beta family of proteins, or inhibits the binding of the TGF-beta binding protein to the TGF-beta family of proteins. Within certain embodiments, the molecule enhances the ability of TGF-beta to function as a positive regulator of mesenchymal cell differentiation.

Within other aspects of the invention, methods are provided for determining whether a selected molecule is capable of increasing bone mineral content, comprising the steps of (a) exposing a selected molecule to cells which express TGF-beta binding-protein and (b) determining whether the expression (or activity) of TGF-beta binding-protein from said exposed cells decreases, and therefrom determining whether the compound is capable of increasing bone mineral content. Within one embodiment, the cells are selected from the group consisting of the spontaneously transformed or untransformed normal human bone from bone biopsies and rat parietal bone osteoblasts. Such methods may be accomplished in a wide variety of assay formats including, for example, Countercurrent Immuno-Electrophoresis (CIEP), Radioimmunoassays, Radioimmunoprecipitations, Enzyme-Linked Immuno-Sorbent Assays (ELISA), Dot Blot assays, Inhibition or Competition assays, and sandwich assays (*see* U.S. Patent Nos. 4,376,110 and 4,486,530; *see also Antibodies: A Laboratory Manual, supra*).

Representative embodiments of such assays are provided below in Examples 5 and 6. Briefly, a family member of the TGF-beta super-family or a TGF-beta binding protein is first bound to a solid phase, followed by addition of a candidate molecule. The labeled family member of the TGF-beta super-family or a TGF-beta binding protein is then added to the assay, the solid phase washed, and the quantity of bound or labeled TGF-beta super-family member or TGF-beta binding protein on the solid support determined. Molecules which are suitable for use in increasing bone mineral content as described herein are those molecules which decrease the binding of TGF-beta binding protein to a member or members of the TGF-beta super-family in a statistically significant manner. Obviously, assays suitable for use within the present invention should not be limited to the embodiments described within Examples 2 and 3. In particular, numerous parameters may be altered, such as by binding TGF-beta to a solid phase, or by elimination of a solid phase entirely.

Within other aspects of the invention, methods are provided for determining whether a selected molecule is capable of increasing bone mineral content, comprising the steps of (a) exposing a selected molecule to cells which express TGF-beta and (b) determining whether the activity of TGF-beta from said exposed cells is altered, and therefrom determining whether the compound is capable of increasing bone mineral content. Similar to the above described methods, a wide variety of methods may be utilized to assess the changes of TGF-beta binding-protein expression due to a selected test compound.



For example, within one aspect of the present invention methods are provided for determining whether a selected molecule is capable of increasing bone mineral content, comprising the steps of (a) mixing a selected molecule with TGF-beta-binding-protein and a selected member of the TGF-beta family of proteins, (b) determining whether the selected molecule up-regulates the signaling of the TGF-beta family of proteins, or inhibits the binding of the TGF-beta binding-protein to the TGF-beta family of proteins. Within certain embodiments, the molecule enhances the ability of TGF-beta to function as a positive regulator of mechamchymal cell differentiation.

Similar to the above described methods, a wide variety of methods may be utilized to assess stimulation of TGF-beta due to a selected test compound. One such representative method is provided below in Example 6 (see also Durham et al., *Endo.* 136:1374-1380.

Within yet other aspects of the present invention, methods are provided for determining whether a selected molecule is capable of increasing bone mineral content, comprising the step of determining whether a selected molecule inhibits the binding of TGF-beta binding-protein to bone, or an analogue thereof. As utilized herein, it should be understood that bone or analogues thereof refers to hydroxyapatite, or a surface composed of a powdered form of bone, crushed bone or intact bone. Similar to the above described methods, a wide variety of methods may be utilized to assess the inhibition of TGF-beta binding-protein localization to bone matrix. One such representative method is provided below in Example 7.

It should be noted that while the methods recited herein may refer to the analysis of an individual test molecule, that the present invention should not be so limited. In particular, the selected molecule may be contained within a mixture of compounds. Hence, the recited methods may further comprise the step of isolating a molecule which inhibits the binding of TGF-beta binding-protein to a TGF-beta family member.

#### CANDIDATE MOLECULES

A wide variety of molecules may be assayed for their ability to inhibit the binding of TGF-beta binding-protein to a TGF-beta family member. Representative examples which are discussed in more detail below include organic molecules, proteins or peptides, and nucleic acid molecules. Although it should be evident from the discussion below that the candidate molecules described herein may be utilized in the

assays described herein, it should also be readily apparent that such molecules can also be utilized in a variety of diagnostic and therapeutic settings.

### 1. Organic Molecules

Numerous organic molecules may be assayed for their ability to inhibit the binding of TGF-beta binding-protein to a TGF-beta family member.

For example, within one embodiment of the invention suitable organic molecules may be selected from either a chemical library, wherein chemicals are assayed individually, or from combinatorial chemical libraries where multiple compounds are assayed at once, then deconvoluted to determine and isolate the most active compounds.

Representative examples of such combinatorial chemical libraries include those described by Agrafiotis et al., "System and method of automatically generating chemical compounds with desired properties," U.S. Patent No. 5,463,564; Armstrong, R.W., "Synthesis of combinatorial arrays of organic compounds through the use of multiple component combinatorial array syntheses," WO 95/02566; Baldwin, J.J. et al., "Sulfonamide derivatives and their use," WO 95/24186; Baldwin, J.J. et al., "Combinatorial dihydrobenzopyran library," WO 95/30642; Brenner, S., "New kit for preparing combinatorial libraries," WO 95/16918; Chenera, B. et al., "Preparation of library of resin-bound aromatic carbocyclic compounds," WO 95/16712; Ellman, J.A., "Solid phase and combinatorial synthesis of benzodiazepine compounds on a solid support," U.S. Patent No. 5,288,514; Felder, E. et al., "Novel combinatorial compound libraries," WO 95/16209; Lerner, R. et al., "Encoded combinatorial chemical libraries," WO 93/20242; Pavia, M.R. et al., "A method for preparing and selecting pharmaceutically useful non-peptide compounds from a structurally diverse universal library," WO 95/04277; Summerton, J.E. and D.D. Weller, "Morpholino-subunit combinatorial library and method," U.S. Patent No. 5,506,337; Holmes, C., "Methods for the Solid Phase Synthesis of Thiazolidinones, Metathiazanones, and Derivatives thereof," WO 96/00148; Phillips, G.B. and G.P. Wei, "Solid-phase Synthesis of Benzimidazoles," *Tet. Letters* 37:4887-90, 1996; Ruhland, B. et al., "Solid-supported Combinatorial Synthesis of Structurally Diverse  $\beta$ -Lactams," *J. Amer. Chem. Soc.* 111:253-4, 1996; Look, G.C. et al., "The Identification of Cyclooxygenase-1 Inhibitors from 4-Thiazolidinone Combinatorial Libraries," *Bioorg and Med. Chem. Letters* 6:707-12, 1996.

## 2. Proteins and Peptides

A wide range of proteins and peptides may likewise be utilized as candidate molecules for inhibitors of the binding of TGF-beta binding-protein to a TGF-beta family member.

### 5 a. Combinatorial Peptide Libraries

Peptide molecules which are putative inhibitors of the binding of TGF-beta binding-protein to a TGF-beta family member may be obtained through the screening of combinatorial peptide libraries. Such libraries may either be prepared by one of skill in the art (*see e.g.*, U.S. Patent Nos. 4,528,266 and 4,359,535, and Patent  
10 Cooperation Treaty Publication Nos. WO 92/15679, WO 92/15677, WO 90/07862, WO 90/02809, or purchased from commercially available sources (*e.g.*, New England Biolabs Ph.D.<sup>TM</sup> Phage Display Peptide Library Kit).

### b. Antibodies

Antibodies which inhibit the binding of TGF-beta binding-protein to a  
15 TGF-beta family member may readily be prepared given the disclosure provided herein. Within the context of the present invention, antibodies are understood to include monoclonal antibodies, polyclonal antibodies, anti-idiotypic antibodies, antibody fragments (*e.g.*, Fab, and F(ab')<sub>2</sub>, F<sub>V</sub> variable regions, or complementarity determining regions). As discussed above, antibodies are understood to be specific  
20 against TGF-beta binding-protein, or against a specific TGF-beta family member, if they bind with a K<sub>a</sub> of greater than or equal to 10<sup>7</sup>M, preferably greater than or equal to 10<sup>8</sup>M, and do not bind to other TGF-beta binding-proteins, or, bind with a K<sub>a</sub> of less than or equal to 10<sup>6</sup>M. Furthermore, antibodies of the present invention should block or inhibit the binding of TGF-beta binding-protein to a TGF-beta family member.

25 The affinity of a monoclonal antibody or binding partner, as well as inhibition of binding can be readily determined by one of ordinary skill in the art (*see* Scatchard, *Ann. N.Y. Acad. Sci.* 51:660-672, 1949).

Briefly, polyclonal antibodies may be readily generated by one of ordinary skill in the art from a variety of warm-blooded animals such as horses, cows,  
30 various fowl, rabbits, mice, or rats. Typically, the TGF-beta binding-protein or unique peptide thereof of 13-20 amino acids (preferably conjugated to keyhole limpet hemocyanin by cross-linking with glutaraldehyde) is utilized to immunize the animal through intraperitoneal, intramuscular, intraocular, or subcutaneous injections, along with an adjuvant such as Freund's complete or incomplete adjuvant. Following several

booster immunizations, samples of serum are collected and tested for reactivity to the protein or peptide. Particularly preferred polyclonal antisera will give a signal on one of these assays that is at least three times greater than background. Once the titer of the animal has reached a plateau in terms of its reactivity to the protein, larger quantities of  
5 antisera may be readily obtained either by weekly bleedings, or by exsanguinating the animal.

Monoclonal antibodies may also be readily generated using conventional techniques (see U.S. Patent Nos. RE 32,011, 4,902,614, 4,543,439, and 4,411,993 which are incorporated herein by reference; see also *Monoclonal Antibodies, Hybridomas: A New Dimension in Biological Analyses*, Plenum Press, Kennett, McKearn, and Bechtol (eds.), 1980, and *Antibodies: A Laboratory Manual*, Harlow and Lane (eds.), Cold Spring Harbor Laboratory Press, 1988, which are also incorporated  
10 herein by reference).

Briefly, within one embodiment a subject animal such as a rat or mouse  
15 is immunized with TGF-beta binding-protein or portion thereof as described above. The protein may be admixed with an adjuvant such as Freund's complete or incomplete adjuvant in order to increase the resultant immune response. Between one and three weeks after the initial immunization the animal may be reimmunized with another booster immunization, and tested for reactivity to the protein utilizing assays described  
20 above. Once the animal has reached a plateau in its reactivity to the injected protein, it is sacrificed, and organs which contain large numbers of B cells such as the spleen and lymph nodes are harvested.

Cells which are obtained from the immunized animal may be immortalized by infection with a virus such as the Epstein-Barr virus (EBV) (see  
25 Glasky and Reading, *Hybridoma* 8(4):377-389, 1989). Alternatively, within a preferred embodiment, the harvested spleen and/or lymph node cell suspensions are fused with a suitable myeloma cell in order to create a "hybridoma" which secretes monoclonal antibody. Suitable myeloma lines include, for example, NS-1 (ATCC No. TIB 18), and P3X63 - Ag 8.653 (ATCC No. CRL 1580).

Following the fusion, the cells may be placed into culture plates  
30 containing a suitable medium, such as RPMI 1640, or DMEM (Dulbecco's Modified Eagles Medium) (JRH Biosciences, Lenexa, Kansas), as well as additional ingredients, such as fetal bovine serum (FBS, *i.e.*, from Hyclone, Logan, Utah, or JRH Biosciences). Additionally, the medium should contain a reagent which selectively  
35 allows for the growth of fused spleen and myeloma cells such as HAT (hypoxanthine, aminopterin, and thymidine) (Sigma Chemical Co., St. Louis, Missouri). After about

seven days, the resulting fused cells or hybridomas may be screened in order to determine the presence of antibodies which are reactive against TGF-beta binding-protein (depending on the antigen used), and which block or inhibit the binding of TGF-beta binding-protein to a TGF-beta family member.

5 A wide variety of assays may be utilized to determine the presence of antibodies which are reactive against the proteins of the present invention, including for example countercurrent immuno-electrophoresis, radioimmunoassays, radioimmunoprecipitations, enzyme-linked immuno-sorbent assays (ELISA), dot blot assays, western blots, immunoprecipitation, inhibition or competition assays, and  
10 sandwich assays (*see* U.S. Patent Nos. 4,376,110 and 4,486,530; *see also* *Antibodies: A Laboratory Manual*, Harlow and Lane (eds.), Cold Spring Harbor Laboratory Press, 1988). Following several clonal dilutions and reassays, a hybridoma producing antibodies reactive against the desired protein may be isolated.

Other techniques may also be utilized to construct monoclonal  
15 antibodies (*see* William D. Huse et al., "Generation of a Large Combinational Library of the Immunoglobulin Repertoire in Phage Lambda," *Science* 246:1275-1281, December 1989; *see also* L. Sastry et al., "Cloning of the Immunological Repertoire in *Escherichia coli* for Generation of Monoclonal Catalytic Antibodies: Construction of a Heavy Chain Variable Region-Specific cDNA Library," *Proc. Natl. Acad. Sci. USA*  
20 86:5728-5732, August 1989; *see also* Michelle Alting-Mees et al., "Monoclonal Antibody Expression Libraries: A Rapid Alternative to Hybridomas," *Strategies in Molecular Biology* 3:1-9, January 1990). These references describe a commercial system available from Stratagene (La Jolla, California) which enables the production of antibodies through recombinant techniques. Briefly, mRNA is isolated from a B cell  
25 population, and utilized to create heavy and light chain immunoglobulin cDNA expression libraries in the  $\lambda$ ImmunoZap(H) and  $\lambda$ ImmunoZap(L) vectors. These vectors may be screened individually or co-expressed to form Fab fragments or antibodies (*see* Huse et al., *supra*; *see also* Sastry et al., *supra*). Positive plaques may subsequently be converted to a non-lytic plasmid which allows high level expression of  
30 monoclonal antibody fragments from *E. coli*.

Similarly, portions or fragments, such as Fab and Fv fragments, of antibodies may also be constructed utilizing conventional enzymatic digestion or recombinant DNA techniques to incorporate the variable regions of a gene which encodes a specifically binding antibody. Within one embodiment, the genes which  
35 encode the variable region from a hybridoma producing a monoclonal antibody of interest are amplified using nucleotide primers for the variable region. These primers

may be synthesized by one of ordinary skill in the art, or may be purchased from commercially available sources. Stratagene (La Jolla, California) sells primers for mouse and human variable regions including, among others, primers for V<sub>Ha</sub>, V<sub>Hb</sub>, V<sub>Hc</sub>, V<sub>Hd</sub>, C<sub>H1</sub>, V<sub>L</sub> and C<sub>L</sub> regions. These primers may be utilized to amplify heavy  
5 or light chain variable regions, which may then be inserted into vectors such as ImmunoZAP™ H or ImmunoZAP™ L (Stratagene), respectively. These vectors may then be introduced into *E. coli*, yeast, or mammalian-based systems for expression. Utilizing these techniques, large amounts of a single-chain protein containing a fusion of the V<sub>H</sub> and V<sub>L</sub> domains may be produced (*see* Bird et al., *Science* 242:423-426,  
10 1988). In addition, such techniques may be utilized to change a "murine" antibody to a "human" antibody, without altering the binding specificity of the antibody.

Once suitable antibodies have been obtained, they may be isolated or purified by many techniques well known to those of ordinary skill in the art (*see* *Antibodies: A Laboratory Manual*, Harlow and Lane (eds.), Cold Spring Harbor  
15 Laboratory Press, 1988). Suitable techniques include peptide or protein affinity columns, HPLC or RP-HPLC, purification on protein A or protein G columns, or any combination of these techniques.

#### c. Mutant TGF-beta binding-proteins

As described herein and below in the Examples (e.g., Examples 8 and  
20 9), altered versions of TGF-beta binding-protein which compete with native TGF-beta binding-protein's ability to block the activity of a particular TGF-beta family member should lead to increased bone density. Thus, mutants of TGF-beta binding-protein which bind to the TGF-beta family member but do not inhibit the function of the TGF-beta family member would meet the criteria. The mutant versions must effectively  
25 compete with the endogenous inhibitory functions of TGF-beta binding-protein.

#### d. Production of proteins

Although various genes (or portions thereof) have been provided herein, it should be understood that within the context of the present invention, reference to one or more of these genes includes derivatives of the genes that are substantially  
30 similar to the genes (and, where appropriate, the proteins (including peptides and polypeptides) that are encoded by the genes and their derivatives). As used herein, a nucleotide sequence is deemed to be "substantially similar" if: (a) the nucleotide sequence is derived from the coding region of the above-described genes and includes, for example, portions of the sequence or allelic variations of the sequences discussed

above, or alternatively, encodes a molecule which inhibits the binding of TGF-beta binding-protein to a member of the TGF-beta family, (b) the nucleotide sequence is capable of hybridization to nucleotide sequences of the present invention under moderate, high or very high stringency (*see* Sambrook et al., *Molecular Cloning: A Laboratory Manual*, 2nd ed., Cold Spring Harbor Laboratory Press, NY, 1989); or (c) the DNA sequences are degenerate as a result of the genetic code to the DNA sequences defined in (a) or (b). Further, the nucleic acid molecule disclosed herein includes both complementary and non-complementary sequences, provided the sequences otherwise meet the criteria set forth herein. Within the context of the present invention, high stringency means standard hybridization conditions (*e.g.*, 5XSSPE, 0.5% SDS at 65°C, or the equivalent).

The structure of the proteins encoded by the nucleic acid molecules described herein may be predicted from the primary translation products using the hydrophobicity plot function of, for example, P/C Gene or Intelligenetics Suite (Intelligenetics, Mountain View, California), or according to the methods described by Kyte and Doolittle (*J. Mol. Biol.* 157:105-132, 1982).

Proteins of the present invention may be prepared in the form of acidic or basic salts, or in neutral form. In addition, individual amino acid residues may be modified by oxidation or reduction. Furthermore, various substitutions, deletions, or additions may be made to the amino acid or nucleic acid sequences, the net effect of which is to retain or further enhance or decrease the biological activity of the mutant or wild-type protein. Moreover, due to degeneracy in the genetic code, for example, there may be considerable variation in nucleotide sequences encoding the same amino acid sequence.

Other derivatives of the proteins disclosed herein include conjugates of the proteins along with other proteins or polypeptides. This may be accomplished, for example, by the synthesis of N-terminal or C-terminal fusion proteins which may be added to facilitate purification or identification of proteins (*see* U.S. Patent No. 4,851,341, *see also*, Hopp et al., *Bio/Technology* 6:1204, 1988.) Alternatively, fusion proteins such as Flag/TGF-beta binding-protein be constructed in order to assist in the identification, expression, and analysis of the protein.

Proteins of the present invention may be constructed using a wide variety of techniques described herein. Further, mutations may be introduced at particular loci by synthesizing oligonucleotides containing a mutant sequence, flanked by restriction sites enabling ligation to fragments of the native sequence. Following

ligation, the resulting reconstructed sequence encodes a derivative having the desired amino acid insertion, substitution, or deletion.

Alternatively, oligonucleotide-directed site-specific (or segment specific) mutagenesis procedures may be employed to provide an altered gene having particular codons altered according to the substitution, deletion, or insertion required. Exemplary methods of making the alterations set forth above are disclosed by Walder et al. (*Gene* 42:133, 1986); Bauer et al. (*Gene* 37:73, 1985); Craik (*BioTechniques*, January 1985, 12-19); Smith et al. (*Genetic Engineering: Principles and Methods*, Plenum Press, 1981); and Sambrook et al. (*supra*). Deletion or truncation derivatives of proteins (*e.g.*, a soluble extracellular portion) may also be constructed by utilizing convenient restriction endonuclease sites adjacent to the desired deletion. Subsequent to restriction, overhangs may be filled in, and the DNA religated. Exemplary methods of making the alterations set forth above are disclosed by Sambrook et al. (*Molecular Cloning: A Laboratory Manual*, 2d Ed., Cold Spring Harbor Laboratory Press, 1989).

Mutations which are made in the nucleic acid molecules of the present invention preferably preserve the reading frame of the coding sequences. Furthermore, the mutations will preferably not create complementary regions that could hybridize to produce secondary mRNA structures, such as loops or hairpins, that would adversely affect translation of the mRNA. Although a mutation site may be predetermined, it is not necessary that the nature of the mutation *per se* be predetermined. For example, in order to select for optimum characteristics of mutants at a given site, random mutagenesis may be conducted at the target codon and the expressed mutants screened for indicative biological activity. Alternatively, mutations may be introduced at particular loci by synthesizing oligonucleotides containing a mutant sequence, flanked by restriction sites enabling ligation to fragments of the native sequence. Following ligation, the resulting reconstructed sequence encodes a derivative having the desired amino acid insertion, substitution, or deletion.

Nucleic acid molecules which encode proteins of the present invention may also be constructed utilizing techniques of PCR mutagenesis, chemical mutagenesis (Drinkwater and Klinedinst, *PNAS* 83:3402-3406, 1986), by forced nucleotide misincorporation (*e.g.*, Liao and Wise *Gene* 88:107-111, 1990), or by use of randomly mutagenized oligonucleotides (Horwitz et al., *Genome* 3:112-117, 1989).

The present invention also provides for the manipulation and expression of the above described genes by culturing host cells containing a vector capable of expressing the above-described genes. Such vectors or vector constructs include either synthetic or cDNA-derived nucleic acid molecules encoding the desired protein, which



are operably linked to suitable transcriptional or translational regulatory elements. Suitable regulatory elements may be derived from a variety of sources, including bacterial, fungal, viral, mammalian, insect, or plant genes. Selection of appropriate regulatory elements is dependent on the host cell chosen, and may be readily  
5 accomplished by one of ordinary skill in the art. Examples of regulatory elements include: a transcriptional promoter and enhancer or RNA polymerase binding sequence, a transcriptional terminator, and a ribosomal binding sequence, including a translation initiation signal.

Nucleic acid molecules that encode any of the proteins described above  
10 may be readily expressed by a wide variety of prokaryotic and eukaryotic host cells, including bacterial, mammalian, yeast or other fungi, viral, insect, or plant cells. Methods for transforming or transfecting such cells to express foreign DNA are well known in the art (see, e.g., Itakura et al., U.S. Patent No. 4,704,362; Hinnen et al.,  
15 *Proc. Natl. Acad. Sci. USA* 75:1929-1933, 1978; Murray et al., U.S. Patent No. 4,801,542; Upshall et al., U.S. Patent No. 4,935,349; Hagen et al., U.S. Patent No. 4,784,950; Axel et al., U.S. Patent No. 4,399,216; Goeddel et al., U.S. Patent No. 4,766,075; and Sambrook et al. *Molecular Cloning: A Laboratory Manual*, 2nd ed., Cold Spring Harbor Laboratory Press, 1989; for plant cells see Czako and Marton,  
20 *Plant Physiol.* 104:1067-1071, 1994; and Paszkowski et al., *Biotech.* 24:387-392, 1992).

Bacterial host cells suitable for carrying out the present invention include *E. coli*, *B. subtilis*, *Salmonella typhimurium*, and various species within the genera *Pseudomonas*, *Streptomyces*, and *Staphylococcus*, as well as many other bacterial species well known to one of ordinary skill in the art. Representative  
25 examples of bacterial host cells include DH5 $\alpha$  (Stratagene, LaJolla, California).

Bacterial expression vectors preferably comprise a promoter which functions in the host cell, one or more selectable phenotypic markers, and a bacterial origin of replication. Representative promoters include the  $\beta$ -lactamase (penicillinase) and lactose promoter system (see Chang et al., *Nature* 275:615, 1978), the T7 RNA  
30 polymerase promoter (Studier et al., *Meth. Enzymol.* 185:60-89, 1990), the lambda promoter (Elvin et al., *Gene* 87:123-126, 1990), the *trp* promoter (Nichols and Yanofsky, *Meth. in Enzymology* 101:155, 1983) and the *tac* promoter (Russell et al., *Gene* 20:231, 1982). Representative selectable markers include various antibiotic resistance markers such as the kanamycin or ampicillin resistance genes. Many  
35 plasmids suitable for transforming host cells are well known in the art, including among others, pBR322 (see Bolivar et al., *Gene* 2:95, 1977), the pUC plasmids pUC18,

pUC19, pUC118, pUC119 (see Messing, *Meth. in Enzymology* 101:20-77, 1983 and Vieira and Messing, *Gene* 19:259-268, 1982), and pNH8A, pNH16a, pNH18a, and Bluescript M13 (Stratagene, La Jolla, California).

Yeast and fungi host cells suitable for carrying out the present invention include, among others, *Saccharomyces pombe*, *Saccharomyces cerevisiae*, the genera *Pichia* or *Kluyveromyces* and various species of the genus *Aspergillus* (McKnight et al., U.S. Patent No. 4,935,349). Suitable expression vectors for yeast and fungi include, among others, YCp50 (ATCC No. 37419) for yeast, and the amdS cloning vector pV3 (Turnbull, *Bio/Technology* 7:169, 1989), YRp7 (Struhl et al., *Proc. Natl. Acad. Sci. USA* 76:1035-1039, 1978), YEp13 (Broach et al., *Gene* 8:121-133, 1979), pJDB249 and pJDB219 (Beggs, *Nature* 275:104-108, 1978) and derivatives thereof.

Preferred promoters for use in yeast include promoters from yeast glycolytic genes (Hitzeman et al., *J. Biol. Chem.* 255:12073-12080, 1980; Alber and Kawasaki, *J. Mol. Appl. Genet.* 1:419-434, 1982) or alcohol dehydrogenase genes (Young et al., in *Genetic Engineering of Microorganisms for Chemicals*, Hollaender et al. (eds.), p. 355, Plenum, New York, 1982; Ammerer, *Meth. Enzymol.* 101:192-201, 1983). Examples of useful promoters for fungi vectors include those derived from *Aspergillus nidulans* glycolytic genes, such as the *adh3* promoter (McKnight et al., *EMBO J.* 4:2093-2099, 1985). The expression units may also include a transcriptional terminator. An example of a suitable terminator is the *adh3* terminator (McKnight et al., *ibid.*, 1985).

As with bacterial vectors, the yeast vectors will generally include a selectable marker, which may be one of any number of genes that exhibit a dominant phenotype for which a phenotypic assay exists to enable transformants to be selected. Preferred selectable markers are those that complement host cell auxotrophy, provide antibiotic resistance or enable a cell to utilize specific carbon sources, and include *leu2* (Broach et al., *ibid.*), *ura3* (Botstein et al., *Gene* 8:17, 1979), or *his3* (Struhl et al., *ibid.*). Another suitable selectable marker is the *cat* gene, which confers chloramphenicol resistance on yeast cells.

Techniques for transforming fungi are well known in the literature, and have been described, for instance, by Beggs (*ibid.*), Hinnen et al. (*Proc. Natl. Acad. Sci. USA* 75:1929-1933, 1978), Yelton et al. (*Proc. Natl. Acad. Sci. USA* 81:1740-1747, 1984), and Russell (*Nature* 301:167-169, 1983). The genotype of the host cell may contain a genetic defect that is complemented by the selectable marker present on the expression vector. Choice of a particular host and selectable marker is well within the level of ordinary skill in the art.

Protocols for the transformation of yeast are also well known to those of ordinary skill in the art. For example, transformation may be readily accomplished either by preparation of spheroplasts of yeast with DNA (*see* Hinnen et al., *PNAS USA* 75:1929, 1978) or by treatment with alkaline salts such as LiCl (*see* Itoh et al., *J. Bacteriology* 153:163, 1983). Transformation of fungi may also be carried out using polyethylene glycol as described by Cullen et al. (*Bio/Technology* 5:369, 1987).

Viral vectors include those which comprise a promoter that directs the expression of an isolated nucleic acid molecule that encodes a desired protein as described above. A wide variety of promoters may be utilized within the context of the present invention, including for example, promoters such as MoMLV LTR, RSV LTR, Friend MuLV LTR, adenoviral promoter (Ohno et al., *Science* 265:781-784, 1994), neomycin phosphotransferase promoter/enhancer, late parvovirus promoter (Koering et al., *Hum. Gene Therap.* 5:457-463, 1994), Herpes TK promoter, SV40 promoter, metallothionein IIa gene enhancer/promoter, cytomegalovirus immediate early promoter, and the cytomegalovirus immediate late promoter. Within particularly preferred embodiments of the invention, the promoter is a tissue-specific promoter (*see e.g.*, WO 91/02805; EP 0,415,731; and WO 90/07936). Representative examples of suitable tissue specific promoters include neural specific enolase promoter, platelet derived growth factor beta promoter, bone morphogenic protein promoter, human alpha1-chimaerin promoter, synapsin I promoter and synapsin II promoter. In addition to the above-noted promoters, other viral-specific promoters (*e.g.*, retroviral promoters (including those noted above, as well as others such as HIV promoters), hepatitis, herpes (*e.g.*, EBV), and bacterial, fungal or parasitic (*e.g.*, malarial) -specific promoters may be utilized in order to target a specific cell or tissue which is infected with a virus, bacteria, fungus or parasite.

Mammalian cells suitable for carrying out the present invention include, among others COS, CHO, SaOS, osteosarcomas, KS483, MG-63, primary osteoblasts, and human or mammalian bone marrow stroma. Mammalian expression vectors for use in carrying out the present invention will include a promoter capable of directing the transcription of a cloned gene or cDNA. Preferred promoters include viral promoters and cellular promoters. Bone specific promoters include the bone sialo-protein and the promoter for osteocalcin. Viral promoters include the cytomegalovirus immediate early promoter (Boshart et al., *Cell* 41:521-530, 1985), cytomegalovirus immediate late promoter, SV40 promoter (Subramani et al., *Mol. Cell. Biol.* 1:854-864, 1981), MMTV LTR, RSV LTR, metallothionein-1, adenovirus E1a. Cellular promoters include the mouse metallothionein-1 promoter (Palmiter et al., U.S. Patent No. 4,579,821), a

mouse V<sub>K</sub> promoter (Bergman et al., *Proc. Natl. Acad. Sci. USA* 81:7041-7045, 1983; Grant et al., *Nucl. Acids Res.* 15:5496, 1987) and a mouse V<sub>H</sub> promoter (Loh et al., *Cell* 33:85-93, 1983). The choice of promoter will depend, at least in part, upon the level of expression desired or the recipient cell line to be transfected.

5 Such expression vectors may also contain a set of RNA splice sites located downstream from the promoter and upstream from the DNA sequence encoding the peptide or protein of interest. Preferred RNA splice sites may be obtained from adenovirus and/or immunoglobulin genes. Also contained in the expression vectors is a polyadenylation signal located downstream of the coding sequence of interest. Suitable  
10 polyadenylation signals include the early or late polyadenylation signals from SV40 (Kaufman and Sharp, *ibid.*), the polyadenylation signal from the Adenovirus 5 E1B region and the human growth hormone gene terminator (DeNoto et al., *Nuc. Acids Res.* 9:3719-3730, 1981). The expression vectors may include a noncoding viral leader sequence, such as the Adenovirus 2 tripartite leader, located between the promoter and  
15 the RNA splice sites. Preferred vectors may also include enhancer sequences, such as the SV40 enhancer. Expression vectors may also include sequences encoding the adenovirus VA RNAs. Suitable expression vectors can be obtained from commercial sources (*e.g.*, Stratagene, La Jolla, California).

Vector constructs comprising cloned DNA sequences can be introduced  
20 into cultured mammalian cells by, for example, calcium phosphate-mediated transfection (Wigler et al., *Cell* 14:725, 1978; Corsaro and Pearson, *Somatic Cell Genetics* 7:603, 1981; Graham and Van der Eb, *Virology* 52:456, 1973), electroporation (Neumann et al., *EMBO J.* 1:841-845, 1982), or DEAE-dextran mediated transfection (Ausubel et al. (eds.), *Current Protocols in Molecular Biology*, John Wiley and Sons,  
25 Inc., NY, 1987). To identify cells that have stably integrated the cloned DNA, a selectable marker is generally introduced into the cells along with the gene or cDNA of interest. Preferred selectable markers for use in cultured mammalian cells include genes that confer resistance to drugs, such as neomycin, hygromycin, and methotrexate. The selectable marker may be an amplifiable selectable marker. Preferred amplifiable  
30 selectable markers are the DHFR gene and the neomycin resistance gene. Selectable markers are reviewed by Thilly (*Mammalian Cell Technology*, Butterworth Publishers, Stoneham, Massachusetts, which is incorporated herein by reference).

Mammalian cells containing a suitable vector are allowed to grow for a period of time, typically 1-2 days, to begin expressing the DNA sequence(s) of interest.  
35 Drug selection is then applied to select for growth of cells that are expressing the selectable marker in a stable fashion. For cells that have been transfected with an

amplifiable, selectable marker the drug concentration may be increased in a stepwise manner to select for increased copy number of the cloned sequences, thereby increasing expression levels. Cells expressing the introduced sequences are selected and screened for production of the protein of interest in the desired form or at the desired level.  
5 Cells that satisfy these criteria can then be cloned and scaled up for production.

Protocols for the transfection of mammalian cells are well known to those of ordinary skill in the art. Representative methods include calcium phosphate mediated transfection, electroporation, lipofection, retroviral, adenoviral and protoplast fusion-mediated transfection (*see* Sambrook et al., *supra*). Naked vector constructs can  
10 also be taken up by muscular cells or other suitable cells subsequent to injection into the muscle of a mammal (or other animals).

Numerous insect host cells known in the art can also be useful within the present invention, in light of the subject specification. For example, the use of baculoviruses as vectors for expressing heterologous DNA sequences in insect cells has  
15 been reviewed by Atkinson et al. (*Pestic. Sci.* 28:215-224, 1990).

Numerous plant host cells known in the art can also be useful within the present invention, in light of the subject specification. For example, the use of *Agrobacterium rhizogenes* as vectors for expressing genes in plant cells has been reviewed by Sinkar et al. (*J. Biosci. (Bangalore)* 11:47-58, 1987).

20 Within related aspects of the present invention, proteins of the present invention may be expressed in a transgenic animal whose germ cells and somatic cells contain a gene which encodes the desired protein and which is operably linked to a promoter effective for the expression of the gene. Alternatively, in a similar manner transgenic animals may be prepared that lack the desired gene (*e.g.*, "knock-out" mice).  
25 Such transgenics may be prepared in a variety of non-human animals, including mice, rats, rabbits, sheep, dogs, goats and pigs (*see* Hammer et al., *Nature* 315:680-683, 1985, Palmiter et al., *Science* 222:809-814, 1983, Brinster et al., *Proc. Natl. Acad. Sci. USA* 82:4438-4442, 1985, Palmiter and Brinster, *Cell* 41:343-345, 1985, and U.S. Patent Nos. 5,175,383, 5,087,571, 4,736,866, 5,387,742, 5,347,075, 5,221,778, and  
30 5,175,384). Briefly, an expression vector, including a nucleic acid molecule to be expressed together with appropriately positioned expression control sequences, is introduced into pronuclei of fertilized eggs, for example, by microinjection. Integration of the injected DNA is detected by blot analysis of DNA from tissue samples. It is preferred that the introduced DNA be incorporated into the germ line of the animal so  
35 that it is passed on to the animal's progeny. Tissue-specific expression may be achieved through the use of a tissue-specific promoter, or through the use of an

inducible promoter, such as the metallothionein gene promoter (Palmiter et al., 1983, *ibid*), which allows regulated expression of the transgene.

Proteins can be isolated by, among other methods, culturing suitable host and vector systems to produce the recombinant translation products of the present invention. Supernatants from such cell lines, or protein inclusions or whole cells where the protein is not excreted into the supernatant, can then be treated by a variety of purification procedures in order to isolate the desired proteins. For example, the supernatant may be first concentrated using commercially available protein concentration filters, such as an Amicon or Millipore Pellicon ultrafiltration unit. Following concentration, the concentrate may be applied to a suitable purification matrix such as, for example, an anti-protein antibody bound to a suitable support. Alternatively, anion or cation exchange resins may be employed in order to purify the protein. As a further alternative, one or more reverse-phase high performance liquid chromatography (RP-HPLC) steps may be employed to further purify the protein. Other methods of isolating the proteins of the present invention are well known in the skill of the art.

A protein is deemed to be "isolated" within the context of the present invention if no other (undesired) protein is detected pursuant to SDS-PAGE analysis followed by Coomassie blue staining. Within other embodiments, the desired protein can be isolated such that no other (undesired) protein is detected pursuant to SDS-PAGE analysis followed by silver staining.

### 3. Nucleic Acid Molecules

Within other aspects of the invention, nucleic acid molecules are provided which are capable of inhibiting TGF-beta binding-protein binding to a member of the TGF-beta family. For example, within one embodiment antisense oligonucleotide molecules are provided which specifically inhibit expression of TGF-beta binding-protein nucleic acid sequences (*see generally*, Hirashima et al. in *Molecular Biology of RNA: New Perspectives* (M. Inouye and B. S. Dudoock, eds., 1987 Academic Press, San Diego, p. 401); *Oligonucleotides: Antisense Inhibitors of Gene Expression* (J.S. Cohen, ed., 1989 MacMillan Press, London); Stein and Cheng, *Science* 261:1004-1012, 1993; WO 95/10607; U.S. Patent No. 5,359,051; WO 92/06693; and EP-A2-612844). Briefly, such molecules are constructed such that they are complementary to, and able to form Watson-Crick base pairs with, a region of transcribed TGF-beta binding-protein mRNA sequence. The resultant double-stranded

nucleic acid interferes with subsequent processing of the mRNA, thereby preventing protein synthesis (see Example 10).

Within other aspects of the invention, ribozymes are provided which are capable of inhibiting the TGF-beta binding-protein binding to a member of the TGF-beta family. As used herein, "ribozymes" are intended to include RNA molecules that contain anti-sense sequences for specific recognition, and an RNA-cleaving enzymatic activity. The catalytic strand cleaves a specific site in a target RNA at greater than stoichiometric concentration. A wide variety of ribozymes may be utilized within the context of the present invention, including for example, the hammerhead ribozyme (for example, as described by Forster and Symons, *Cell* 48:211-220, 1987; Haseloff and Gerlach, *Nature* 328:596-600, 1988; Walbot and Bruening, *Nature* 334:196, 1988; Haseloff and Gerlach, *Nature* 334:535, 1988); the hairpin ribozyme (for example, as described by Haseloff et al., U.S. Patent No. 5,254,678, issued October 19, 1993 and Hempel et al., European Patent Publication No. 0 360 257, published March 26, 1990); and *Tetrahymena* ribosomal RNA-based ribozymes (see Cech et al., U.S. Patent No. 4,987,071). Ribozymes of the present invention typically consist of RNA, but may also be composed of DNA, nucleic acid analogs (e.g., phosphorothioates), or chimerics thereof (e.g., DNA/RNA/RNA).

#### 4. Labels

The gene product or any of the candidate molecules described above and below, may be labeled with a variety of compounds, including for example, fluorescent molecules, toxins, and radionuclides. Representative examples of fluorescent molecules include fluorescein, *Phycobili* proteins, such as phycoerythrin, rhodamine, Texas red and luciferase. Representative examples of toxins include ricin, abrin diphtheria toxin, cholera toxin, gelonin, pokeweed antiviral protein, tritin, *Shigella* toxin, and *Pseudomonas* exotoxin A. Representative examples of radionuclides include Cu-64, Ga-67, Ga-68, Zr-89, Ru-97, Tc-99m, Rh-105, Pd-109, In-111, I-123, I-125, I-131, Re-186, Re-188, Au-198, Au-199, Pb-203, At-211, Pb-212 and Bi-212. In addition, the antibodies described above may also be labeled or conjugated to one partner of a ligand binding pair. Representative examples include avidin-biotin, and riboflavin-riboflavin binding protein.

Methods for conjugating or labeling the molecules described herein with the representative labels set forth above may be readily accomplished by one of ordinary skill in the art (see Trichothecene Antibody Conjugate, U.S. Patent No. 4,744,981; Antibody Conjugate, U.S. Patent No. 5,106,951; Fluorogenic Materials and

Labeling Techniques, U.S. Patent No. 4,018,884; Metal Radionuclide Labeled Proteins for Diagnosis and Therapy, U.S. Patent No. 4,897,255; and Metal Radionuclide Chelating Compounds for Improved Chelation Kinetics, U.S. Patent No. 4,988,496; *see also* Inman, *Methods In Enzymology*, Vol. 34, *Affinity Techniques, Enzyme*  
5 *Purification: Part B*, Jakoby and Wilchek (eds.), Academic Press, New York, p. 30, 1974; *see also* Wilchek and Bayer, "The Avidin-Biotin Complex in Bioanalytical Applications," *Anal. Biochem.* 171:1-32, 1988).

#### PHARMACEUTICAL COMPOSITIONS

As noted above, the present invention also provides a variety of  
10 pharmaceutical compositions, comprising one of the above-described molecules which inhibits the TGF-beta binding-protein binding to a member of the TGF-beta family along with a pharmaceutically or physiologically acceptable carrier, excipients or diluents. Generally, such carriers should be nontoxic to recipients at the dosages and concentrations employed. Ordinarily, the preparation of such compositions entails  
15 combining the therapeutic agent with buffers, antioxidants such as ascorbic acid, low molecular weight (less than about 10 residues) polypeptides, proteins, amino acids, carbohydrates including glucose, sucrose or dextrans, chelating agents such as EDTA, glutathione and other stabilizers and excipients. Neutral buffered saline or saline mixed with nonspecific serum albumin are exemplary appropriate diluents.

20 In addition, the pharmaceutical compositions of the present invention may be prepared for administration by a variety of different routes. In addition, pharmaceutical compositions of the present invention may be placed within containers, along with packaging material which provides instructions regarding the use of such pharmaceutical compositions. Generally, such instructions will include a tangible  
25 expression describing the reagent concentration, as well as within certain embodiments, relative amounts of excipient ingredients or diluents (*e.g.*, water, saline or PBS) which may be necessary to reconstitute the pharmaceutical composition.

#### METHODS OF TREATMENT

The present invention also provides methods for increasing the mineral  
30 content and mineral density of bone. Briefly, numerous conditions result in the loss of bone mineral content, including for example, disease, genetic predisposition, accidents which result in the lack of use of bone (*e.g.*, due to fracture), therapeutics which effect bone resorption, or which kill bone forming cells and normal aging. Through use of the molecules described herein which inhibit the TGF-beta binding-protein binding to a



TGF-beta family member such conditions may be treated or prevented. As utilized herein, it should be understood that bone mineral content has been increased, if bone mineral content has been increased in a statistically significant manner (*e.g.*, greater than one-half standard deviation), at a selected site.

5 A wide variety of conditions which result in loss of bone mineral content may be treated with the molecules described herein. Patients with such conditions may be identified through clinical diagnosis utilizing well known techniques (see, *e.g.*, Harrison's Principles of Internal Medicine, McGraw-Hill, Inc.). Representative examples of diseases that may be treated included dysplasias, wherein there is abnormal  
10 growth or development of bone. Representative examples of such conditions include achondroplasia, cleidocranial dysostosis, enchondromatosis, fibrous dysplasia, Gaucher's, hypophosphatemic rickets, Marfan's, multiple hereditary exotoses, neurofibromatosis, osteogenesis imperfecta, osteopetrosis, osteopoikilosis, sclerotic lesions, fractures, periodontal disease, pseudoarthrosis and pyogenic osteomyelitis.

15 Other conditions which may be treated or prevented include a wide variety of causes of osteopenia (*i.e.*, a condition that causes greater than one standard deviation of bone mineral content or density below peak skeletal mineral content at youth). Representative examples of such conditions include anemic states, conditions caused steroids, conditions caused by heparin, bone marrow disorders, scurvy,  
20 malnutrition, calcium deficiency, idiopathic osteoporosis, congenital osteopenia or osteoporosis, alcoholism, chronic liver disease, senility, postmenopausal state, oligomenorrhea, amenorrhea, pregnancy, diabetes mellitus, hyperthyroidism, Cushing's disease, acromegaly, hypogonadism, immobilization or disuse, reflex sympathetic dystrophy syndrome, transient regional osteoporosis and osteomalacia.

25 Within one aspect of the present invention, bone mineral content or density may be increased by administering to a warm-blooded animal a therapeutically effective amount of a molecule which inhibits the TGF-beta binding-protein binding to a TGF-beta family member. Examples of warm-blooded animals that may be treated include both vertebrates and mammals, including for example horses, cows, pigs,  
30 sheep, dogs, cats, rats and mice. Representative examples of therapeutic molecules include ribozymes, ribozyme genes, antisense oligonucleotides and antibodies (*e.g.*, humanized antibodies).

Within other aspects of the present invention, methods are provided for increasing bone density, comprising the step of introducing into cells which home to  
35 bone a vector which directs the expression of a molecule which inhibits the TGF-beta binding-protein binding to a member of the TGF-beta family, and administering the

vector containing cells to a warm-blooded animal. Briefly, cells which home to bone may be obtained directly from the bone of patients (e.g., cells obtained from the bone marrow such as CD34+, osteoblasts, osteocytes, and the like), from peripheral blood, or from cultures.

5 A vector which directs the expression of a molecule that inhibits the TGF-beta binding-protein binding to a member of the TGF-beta family is introduced into the cells. Representative examples of suitable vectors include viral vectors such as herpes viral vectors (e.g., U.S. Patent No. 5,288,641), adenoviral vectors (e.g., WO 94/26914, WO 93/9191; Kolls et al., *PNAS* 91(1):215-219, 1994; Kass-Eisler et al.,  
10 *PNAS* 90(24):11498-502, 1993; Guzman et al., *Circulation* 88(6):2838-48, 1993; Guzman et al., *Cir. Res.* 73(6):1202-1207, 1993; Zabner et al., *Cell* 75(2):207-216, 1993; Li et al., *Hum Gene Ther.* 4(4):403-409, 1993; Caillaud et al., *Eur. J. Neurosci.* 5(10):1287-1291, 1993; Vincent et al., *Nat. Genet.* 5(2):130-134, 1993; Jaffe et al., *Nat. Genet.* 1(5):372-378, 1992; and Levrero et al., *Gene* 101(2):195-202, 1991), adeno-  
15 associated viral vectors (WO 95/13365; Flotte et al., *PNAS* 90(22):10613-10617, 1993), baculovirus vectors, parvovirus vectors (Koering et al., *Hum. Gene Therap.* 5:457-463, 1994), pox virus vectors (Panicali and Paoletti, *PNAS* 79:4927-4931, 1982; and Ozaki et al., *Biochem. Biophys. Res. Comm.* 193(2):653-660, 1993), and retroviruses (e.g., EP 0,415,731; WO 90/07936; WO 91/0285, WO 94/03622; WO 93/25698; WO 93/25234; U.S. Patent No. 5,219,740; WO 93/11230; WO 93/10218). Viral  
20 vectors may likewise be constructed which contain a mixture of different elements (e.g., promoters, envelope sequences and the like) from different viruses, or non-viral sources. Within various embodiments, either the viral vector itself, or a viral particle which contains the viral vector may be utilized in the methods and compositions  
25 described below.

Within other embodiments of the invention, nucleic acid molecules which encode a molecule which inhibits the TGF-beta binding-protein binding to a member of the TGF-beta family themselves may be administered by a variety of techniques, including, for example, administration of asialoosomucoid (ASOR)  
30 conjugated with poly-L-lysine DNA complexes (Cristano et al., *PNAS* 92:122-92126, 1993), DNA linked to killed adenovirus (Curiel et al., *Hum. Gene Ther.* 3(2):147-154, 1992), cytofectin-mediated introduction (DMRIE-DOPE, Vical, California), direct DNA injection (Acsadi et al., *Nature* 352:815-818, 1991); DNA ligand (Wu et al., *J. of Biol. Chem.* 264:16985-16987, 1989); lipofection (Felgner et al., *Proc. Natl. Acad. Sci. USA* 84:7413-7417, 1989); liposomes (Pickering et al., *Circ.* 89(1):13-21, 1994; and  
35 Wang et al., *PNAS* 84:7851-7855, 1987); microprojectile bombardment (Williams

et al., *PNAS* 88:2726-2730, 1991); and direct delivery of nucleic acids which encode the protein itself either alone (Vile and Hart, *Cancer Res.* 53: 3860-3864, 1993), or utilizing PEG-nucleic acid complexes.

Representative examples of molecules which may be expressed by the  
5 vectors of present invention include ribozymes and antisense molecules, each of which are discussed in more detail above.

Determination of increased bone mineral content may be determined directly through the use of X-rays (e.g., Dual Energy X-ray Absorptometry or "DEXA"), or by inference through bone turnover markers (osteoblast specific alkaline  
10 phosphatase, osteocalcin, type 1 procollagen C' propeptide (PICP), and total alkaline phosphatase; see Comier, C., *Curr. Opin. in Rheu.* 7:243, 1995), or markers of bone resorption (pyridinoline, deoxypyridinoline, N-telopeptide, urinary hydroxyproline, plasma tartrate-resistant acid phosphatases and galactosyl hydroxylysine; see Comier,  
15 *supra*). The amount of bone mass may also be calculated from body weights, or utilizing other methods (see Guinness-Hey, *Metab. Bone Dis. and Rel. Res.* 5:177-181, 1984).

As will be evident to one of skill in the art, the amount and frequency of administration will depend, of course, on such factors as the nature and severity of the indication being treated, the desired response, the condition of the patient, and so forth.  
20 Typically, the compositions may be administered by a variety of techniques, as noted above.

The following examples are offered by way of illustration, and not by way of limitation.

## EXAMPLES

EXAMPLE 1

## SCLEROSTEOSIS MAPS TO THE LONG ARM OF HUMAN CHROMOSOME 17

Genetic mapping of the defect responsible for sclerosteosis in humans localized the gene responsible for this disorder to the region of human chromosome 17 that encodes a novel TGF-beta binding-protein family member. In sclerosteosis, skeletal bone displays a substantial increase in mineral density relative to that of unaffected individuals. Bone in the head displays overgrowth as well. Sclerosteosis patients are generally healthy although they may exhibit variable degrees of syndactyly at birth and variable degrees of cranial compression and nerve compression in the skull.

Linkage analysis of the gene defect associated with sclerosteosis was conducted by applying the homozygosity mapping method to DNA samples collected from 24 South African Afrikaaner families in which the disease occurred. (Sheffield et al., 1994, *Human Molecular Genetics* 3:1331-1335. "Identification of a Bardet-Biedl syndrome locus on chromosome 3 and evaluation of an efficient approach to homozygosity mapping"). The Afrikaaner population of South Africa is genetically homogeneous; the population is descended from a small number of founders who colonized the area several centuries ago, and it has been isolated by geographic and social barriers since the founding. Sclerosteosis is rare everywhere in the world outside the Afrikaaner community, which suggests that a mutation in the gene was present in the founding population and has since increased in numbers along with the increase in the population. The use of homozygosity mapping is based on the assumption that DNA mapping markers adjacent to a recessive mutation are likely to be homozygous in affected individuals from consanguineous families and isolated populations.

A set of 371 microsatellite markers (Research Genetics, Set 6) from the autosomal chromosomes was selected to type pools of DNA from sclerosteosis patient samples. The DNA samples for this analysis came from 29 sclerosteosis patients in 24 families, 59 unaffected family members and a set of unrelated control individuals from the same population. The pools consisted of 4-6 individuals, either affected individuals, affected individuals from consanguineous families, parents and unaffected siblings, or unrelated controls. In the pools of unrelated individuals and in most of the pools with affected individuals or family members analysis of the markers showed several allele sizes for each marker. One marker, D17S1299, showed an indication of homozygosity: one band in several of the pools of affected individuals.

All 24 sclerosteosis families were typed with a total of 19 markers in the region of D17S1299 (at 17q12-q21). Affected individuals from every family were shown to be homozygous in this region, and 25 of the 29 individuals were homozygous for a core haplotype; they each had the same alleles between D17S1787 and D17S930. The other four individuals had one chromosome which matched this haplotype and a second which did not. In sum, the data compellingly suggested that this 3 megabase region contained the sclerosteosis mutation. Sequence analysis of most of the exons in this 3 megabase region identified a nonsense mutation in the novel TGF-beta binding-protein coding sequence (C>T mutation at position 117 of Sequence ID No. 1 results in a stop codon). This mutation was shown to be unique to sclerosteosis patients and carriers of Afrikaaner descent. The identity of the gene was further confirmed by identifying a mutation in its intron (A>T mutation at position +3 of the intron) which results in improper mRNA processing in a single, unrelated patient with diagnosed sclerosteosis.

## EXAMPLE 2

### TISSUE-SPECIFICITY OF TGF-BETA BINDING-PROTEIN GENE EXPRESSION

#### A. Human Beer Gene Expression by RT-PCR:

First-strand cDNA was prepared from the following total RNA samples using a commercially available kit ("Superscript Preamplification System for First-Strand cDNA Synthesis", Life Technologies, Rockville, MD): human brain, human liver, human spleen, human thymus, human placenta, human skeletal muscle, human thyroid, human pituitary, human osteoblast (NHOst from Clonetics Corp., San Diego, CA), human osteosarcoma cell line (Saos-2, ATCC# HTB-85), human bone, human bone marrow, human cartilage, vervet monkey bone, saccharomyces cerevisiae, and human peripheral blood monocytes. All RNA samples were purchased from a commercial source (Clontech, Palo Alto, CA), except the following which were prepared in-house: human osteoblast, human osteosarcoma cell line, human bone, human cartilage and vervet monkey bone. These in-house RNA samples were prepared using a commercially available kit ("TRI Reagent", Molecular Research Center, Inc., Cincinnati, OH).

PCR was performed on these samples, and additionally on a human genomic sample as a control. The sense Beer oligonucleotide primer had the sequence 5'-CCGGAGCTGGAGAACAACAAG-3' (SEQ ID NO:19). The antisense Beer oligonucleotide primer had the sequence 5'-GCACTGGCCGGAGCACACC-3' (SEQ

ID NO:20). In addition, PCR was performed using primers for the human beta-actin gene, as a control. The sense beta-actin oligonucleotide primer had the sequence 5'-AGGCCAACCGCGAGAAGATGA CC -3' (SEQ ID NO:21). The antisense beta-actin oligonucleotide primer had the sequence 5'-GAAGT CCAGGGCGACGTAGCA-3' (SEQ ID NO:22). PCR was performed using standard conditions in 25 ul reactions, with an annealing temperature of 61 degrees Celsius. Thirty-two cycles of PCR were performed with the Beer primers and twenty-four cycles were performed with the beta-actin primers.

Following amplification, 12 ul from each reaction were analyzed by agarose gel electrophoresis and ethidium bromide staining. See Figure 2A.

#### B. RNA In-situ Hybridization of Mouse Embryo Sections:

The full length mouse *Beer* cDNA (Sequence ID No. 11) was cloned into the pCR2.1 vector (Invitrogen, Carlsbad, CA) in the antisense and sense direction using the manufacturer's protocol. <sup>35</sup>S-alpha-GTP-labeled cRNA sense and antisense transcripts were synthesized using in-vitro transcription reagents supplied by Ambion, Inc (Austin, TX). In-situ hybridization was performed according to the protocols of Lyons et al. (*J. Cell Biol.* 111:2427-2436, 1990).

The mouse *Beer* cRNA probe detected a specific message expressed in the neural tube, limb buds, blood vessels and ossifying cartilages of developing mouse embryos. Panel A in Figure 3 shows expression in the apical ectodermal ridge (aer) of the limb (l) bud, blood vessels (bv) and the neural tube (nt). Panel B shows expression in the 4<sup>th</sup> ventricle of the brain (4). Panel C shows expression in the mandible (ma) cervical vertebrae (cv), occipital bone (oc), palate (pa) and a blood vessel (bv). Panel D shows expression in the ribs (r) and a heart valve (va). Panel A is a transverse section of 10.5 dpc embryo. Panel B is a sagittal section of 12.5 dpc embryo and panels C and D are sagittal sections of 15.5 dpc embryos.

ba=branchial arch, h=heart, te=telencephalon (forebrain), b=brain, f=frontonasal mass, g=gut, h=heart, j=jaw, li=liver, lu=lung, ot=otic vesicle, ao=, sc=spinal cord, skm=skeletal muscle, ns=nasal sinus, th=thymus, to=tongue, fl=forelimb, di=diaphragm

### EXAMPLE 3

#### 35 EXPRESSION AND PURIFICATION OF RECOMBINANT BEER PROTEIN

##### A. Expression in COS-1 Cells:

The DNA sequence encoding the full length human Beer protein was amplified using the following PCR oligonucleotide primers: The 5' oligonucleotide primer had the sequence 5'-**AAGCTTGGTACCATGCAGCTCCAC**-3' (SEQ ID NO:23) and contained a HindIII restriction enzyme site (in bold) followed by 19 nucleotides of the *Beer* gene starting 6 base pairs prior to the presumed amino terminal start codon (ATG). The 3' oligonucleotide primer had the sequence 5'-**AAGCTTCTACTTGTTCATCGTCGTCCT** TGTAGTCGTAGGCGTTCTCCAGCT-3' (SEQ ID NO:24) and contained a HindIII restriction enzyme site (in bold) followed by a reverse complement stop codon (CTA) followed by the reverse complement of the FLAG epitope (underlined, Sigma-Aldrich Co., St. Louis, MO) flanked by the reverse complement of nucleotides coding for the carboxy terminal 5 amino acids of the Beer. The PCR product was TA cloned ("Original TA Cloning Kit", Invitrogen, Carlsbad, CA) and individual clones were screened by DNA sequencing. A sequence-verified clone was then digested by HindIII and purified on a 1.5% agarose gel using a commercially available reagents ("QIAquick Gel Extraction Kit", Qiagen Inc., Valencia, CA). This fragment was then ligated to HindIII digested, phosphatase-treated pcDNA3.1 (Invitrogen, Carlsbad, CA) plasmid with T4 DNA ligase. DH10B *E. coli* were transformed and plated on LB, 100 µg/ml ampicillin plates. Colonies bearing the desired recombinant in the proper orientation were identified by a PCR-based screen, using a 5' primer corresponding to the T7 promoter/priming site in pcDNA3.1 and a 3' primer with the sequence 5'- GCACTGGCCGGAGCACACC-3' (SEQ ID NO:25) that corresponds to the reverse complement of internal BEER sequence. The sequence of the cloned fragment was confirmed by DNA sequencing.

COS-1 cells (ATCC# CRL-1650) were used for transfection. 50 µg of the expression plasmid pcDNA-Beer-Flag was transfected using a commercially available kit following protocols supplied by the manufacturer ("DEAE-Dextran Transfection Kit", Sigma Chemical Co., St. Louis, MO). The final media following transfection was DMEM (Life Technologies, Rockville, MD) containing 0.1% Fetal Bovine Serum. After 4 days in culture, the media was removed. Expression of recombinant BEER was analyzed by SDS-PAGE and Western Blot using anti-FLAG M2 monoclonal antibody (Sigma-Aldrich Co., St. Louis, MO). Purification of recombinant BEER protein was performed using an anti-FLAG M2 affinity column ("Mammalian Transient Expression System", Sigma-Aldrich Co., St. Louis, MO). The column profile was analyzed via SDS-PAGE and Western Blot using anti-FLAG M2 monoclonal antibody.

B. Expression in SF9 insect cells:

The human *Beer* gene sequence was amplified using PCR with standard conditions and the following primers:

Sense primer: 5'-GTCGTCGGATCCATGGGGTGGCAGGCGTTCAAGAATGAT-3'  
5 (SEQ ID NO:26)

Antisense primer: 5'-GTCGTCAAGCTTCTACTTGTTCATCGTCCTTGTAGTCGTA  
GGCGTTCTCCAGCTCGGC-3' (SEQ ID NO:27)

The resulting cDNA contained the coding region of Beer with two modifications. The N-terminal secretion signal was removed and a FLAG epitope tag  
10 (Sigma) was fused in frame to the C-terminal end of the insert. BamHI and HindIII cloning sites were added and the gene was subcloned into pMelBac vector (Invitrogen) for transfer into a baculoviral expression vector using standard methods.

Recombinant baculoviruses expressing Beer protein were made using the Bac-N-Blue transfection kit (Invitrogen) and purified according to the manufacturers  
15 instructions.

SF9 cells (Invitrogen) were maintained in TNM\_FH media (Invitrogen) containing 10% fetal calf serum. For protein expression, SF9 cultures in spinner flasks were infected at an MOI of greater than 10. Samples of the media and cells were taken daily for five days, and Beer expression monitored by western blot using an anti-FLAG  
20 M2 monoclonal antibody (Sigma) or an anti-Beer rabbit polyclonal antiserum.

After five days the baculovirus-infected SF9 cells were harvested by centrifugation and cell associated protein was extracted from the cell pellet using a high salt extraction buffer (1.5 M NaCl, 50 mM Tris pH 7.5). The extract (20 ml per 300 ml culture) was clarified by centrifugation, dialyzed three times against four liters of  
25 Tris buffered saline (150 mM NaCl, 50 mM Tris pH 7.5), and clarified by centrifugation again. This high salt fraction was applied to Hitrap Heparin (Pharmacia; 5 ml bed volume), washed extensively with HEPES buffered saline (25 mM HEPES 7.5, 150 mM NaCl) and bound proteins were eluted with a gradient from 150 mM NaCl to 1200 mM NaCl. Beer elution was observed at approximately 800 mM NaCl. Beer  
30 containing fractions were supplemented to 10% glycerol and 1 mM DTT and frozen at -80 degrees C.

EXAMPLE 4

35 PREPARATION AND TESTING OF POLYCLONAL ANTIBODIES TO BEER, GREMLIN, AND  
DAN



A. Preparation of antigen:

The DNA sequences of Human *Beer*, Human *Gremlin*, and Human *Dan* were amplified using standard PCR methods with the following oligonucleotide primers:

5 H. Beer

Sense: 5' -GACTTGGATCCCAGGGGTGGCAGGCGTTC- 3' (SEQ ID NO:28)

Antisense 5' -AGCATAAGCTTCTAGTAGGCGTTCTCCAG- 3' (SEQ ID NO:29)

H. Gremlin

Sense: 5' -GACTTGGATCCGAAGGGGAAAAAGAAAGGG- 3' (SEQ ID NO:30)

10 Antisense: 5' -AGCATAAGCTTTTAATCCAAATCGATGGA- 3' (SEQ ID NO:31)

H. Dan

Sense: 5' -ACTACGAGCTCGGCCCCACCCATCAACAAG- 3' (SEQ ID NO:32)

Antisense: 5' -ACTTAGAAGCTTTCAGTCCTCAGCCCCCTCTTCC-3' (SEQ ID NO:33)

15 In each case the listed primers amplified the entire coding region minus the secretion signal sequence. These include restriction sites for subcloning into the bacterial expression vector pQE-30 (Qiagen Inc., Valencia, CA) at sites BamHI/HindIII for Beer and Gremlin, and sites SacI/HindIII for Dan. pQE30 contains a coding sequence for a 6x His tag at the 5' end of the cloning region. The completed constructs  
20 were transformed into *E. coli* strain M-15/pRep (Qiagen Inc) and individual clones verified by sequencing. Protein expression in M-15/pRep and purification (6xHis affinity tag binding to Ni-NTA coupled to Sepharose) were performed as described by the manufacturer (Qiagen, The QIAexpressionist).

The *E. coli*-derived Beer protein was recovered in significant quantity  
25 using solubilization in 6M guanidine and dialyzed to 2-4M to prevent precipitation during storage. Gremlin and Dan protein were recovered in higher quantity with solubilization in 6M guanidine and a post purification guanidine concentration of 0.5M.

B. Production and testing of polyclonal antibodies:

30 Polyclonal antibodies to each of the three antigens were produced in rabbit and in chicken hosts using standard protocols (R & R Antibody, Stanwood, WA; standard protocol for rabbit immunization and antisera recovery; Short Protocols in Molecular Biology. 2nd edition. 1992. 11.37- 11.41. Contributors Helen M. Cooper and Yvonne Paterson; chicken antisera was generated with Strategic Biosolutions,  
35 Ramona, CA).

Rabbit antisera and chicken egg IgY fraction were screened for activity

via Western blot. Each of the three antigens was separated by PAGE and transferred to 0.45um nitrocellulose (Novex, San Diego, CA). The membrane was cut into strips with each strip containing approximately 75 ng of antigen. The strips were blocked in 3% Blotting Grade Block (Bio-Rad Laboratories, Hercules, CA) and washed 3 times in 1X Tris buffer saline (TBS) /0.02% TWEEN buffer. The primary antibody (preimmunization bleeds, rabbit antisera or chicken egg IgY in dilutions ranging from 1:100 to 1:10,000 in blocking buffer) was incubated with the strips for one hour with gentle rocking. A second series of three washes 1X TBS/0.02%TWEEN was followed by an one hour incubation with the secondary antibody (peroxidase conjugated donkey anti-rabbit, Amersham Life Science, Piscataway, NJ; or peroxidase conjugated donkey anti-chicken, Jackson ImmunoResearch, West Grove, PA). A final cycle of 3X washes of 1X TBS/0.02%TWEEN was performed and the strips were developed with Lumi-Light Western Blotting Substrate (Roche Molecular Biochemicals, Mannheim, Germany).

15

C. Antibody cross-reactivity test:

Following the protocol described in the previous section, nitrocellulose strips of Beer, Gremlin or Dan were incubated with dilutions (1:5000 and 1:10,000) of their respective rabbit antisera or chicken egg IgY as well as to antisera or chicken egg Igy (dilutions 1:1000 and 1:5000) made to the remaining two antigens. The increased levels of nonmatching antibodies was performed to detect low affinity binding by those antibodies that may be seen only at increased concentration. The protocol and duration of development is the same for all three binding events using the protocol described above. There was no antigen cross-reactivity observed for any of the antigens tested.

25

### EXAMPLE 5

#### INTERACTION OF BEER WITH TGF-BETA SUPER-FAMILY PROTEINS

The interaction of Beer with proteins from different phylogenetic arms of the TGF- $\beta$  superfamily were studied using immunoprecipitation methods. Purified TGF $\beta$ -1, TGF $\beta$ -2, TGF $\beta$ -3, BMP-4, BMP-5, BMP-6 and GDNF were obtained from commercial sources (R&D systems; Minneapolis, MN). A representative protocol is as follows. Partially purified Beer was dialyzed into HEPES buffered saline (25 mM HEPES 7.5, 150 mM NaCl). Immunoprecipitations were done in 300 ul of IP buffer (150 mM NaCl, 25 mM Tris pH 7.5, 1mM EDTA, 1.4 mM  $\beta$ -mercaptoethanol, 0.5 % triton X 100, and 10% glycerol). 30 ng recombinant human BMP-5 protein (R&D

systems) was applied to 15  $\mu$ l of FLAG affinity matrix (Sigma; St Louis MO)) in the presence and absence of 500 ng FLAG epitope-tagged Beer. The proteins were incubated for 4 hours @ 4°C and then the affinity matrix-associated proteins were washed 5 times in IP buffer (1 ml per wash). The bound proteins were eluted from the affinity matrix in 60 microliters of 1X SDS PAGE sample buffer. The proteins were resolved by SDS PAGE and Beer associated BMP-5 was detected by western blot using anti-BMP-5 antiserum (Research Diagnostics, Inc) (see Figure 5).

#### BEER Ligand Binding Assay:

FLAG-Beer protein (20 ng) is added to 100  $\mu$ l PBS/0.2% BSA and adsorbed into each well of 96 well microtiter plate previously coated with anti-FLAG monoclonal antibody (Sigma; St Louis MO) and blocked with 10% BSA in PBS. This is conducted at room temperature for 60 minutes. This protein solution is removed and the wells are washed to remove unbound protein. BMP-5 is added to each well in concentrations ranging from 10 pM to 500 nM in PBS/0.2% BSA and incubated for 2 hours at room temperature. The binding solution is removed and the plate washed with three times with 200  $\mu$ l volumes of PBS/0.2% BSA. BMP-5 levels are then detected using BMP-5 anti-serum via ELISA (F.M. Ausubel et al (1998) Current Protocols in Mol Biol. Vol 2 11.2.1-11.2.22). Specific binding is calculated by subtracting non-specific binding from total binding and analyzed by the LIGAND program (Munson and Podbard, Anal. Biochem., 107, p220-239, (1980)).

In a variation of this method, Beer is engineered and expressed as a human Fc fusion protein. Likewise the ligand BMP is engineered and expressed as mouse Fc fusion. These proteins are incubated together and the assay conducted as described by Mellor et al using homogeneous time resolved fluorescence detection (G.W. Mellor et al., *J of Biomol Screening*, 3(2) 91-99, 1998).

#### EXAMPLE 6

##### SCREENING ASSAY FOR INHIBITION OF TGF-BETA BINDING-PROTEIN BINDING TO TGF-BETA FAMILY MEMBERS

The assay described above is replicated with two exceptions. First, BMP concentration is held fixed at the  $K_d$  determined previously. Second, a collection of antagonist candidates is added at a fixed concentration (20  $\mu$ M in the case of the small organic molecule collections and 1  $\mu$ M in antibody studies). These candidate molecules (antagonists) of TGF-beta binding-protein binding include organic

compounds derived from commercial or internal collections representing diverse chemical structures. These compounds are prepared as stock solutions in DMSO and are added to assay wells at  $\leq 1\%$  of final volume under the standard assay conditions. These are incubated for 2 hours at room temperature with the BMP and Beer, the solution removed and the bound BMP is quantitated as described. Agents that inhibit 40% of the BMP binding observed in the absence of compound or antibody are considered antagonists of this interaction. These are further evaluated as potential inhibitors based on titration studies to determine their inhibition constants and their influence on TGF-beta binding-protein binding affinity. Comparable specificity control assays may also be conducted to establish the selectivity profile for the identified antagonist through studies using assays dependent on the BMP ligand action (e.g. BMP/BMP receptor competition study).

#### EXAMPLE 7

##### INHIBITION OF TGF-BETA BINDING-PROTEIN LOCALIZATION TO BONE MATRIX

Evaluation of inhibition of localization to bone matrix (hydroxyapatite) is conducted using modifications to the method of Nicolas (Nicolas, V. *Calcif Tissue Int* 57:206, 1995). Briefly,  $^{125}\text{I}$ -labelled TGF-beta binding-protein is prepared as described by Nicolas (*supra*). Hydroxyapatite is added to each well of a 96 well microtiter plate equipped with a polypropylene filtration membrane (Polyfiltroninc, Weymouth MA). TGF-beta binding-protein is added to 0.2% albumin in PBS buffer. The wells containing matrix are washed 3 times with this buffer. Adsorbed TGF-beta binding-protein is eluted using 0.3M NaOH and quantitated.

Inhibitor identification is conducted via incubation of TGF-beta binding-protein with test molecules and applying the mixture to the matrix as described above. The matrix is washed 3 times with 0.2% albumin in PBS buffer. Adsorbed TGF-beta binding-protein is eluted using 0.3 M NaOH and quantitated. Agents that inhibit 40% of the TGF-beta binding-protein binding observed in the absence of compound or antibody are considered bone localization inhibitors. These inhibitors are further characterized through dose response studies to determine their inhibition constants and their influence on TGF-beta binding-protein binding affinity.

### EXAMPLE 8

#### CONSTRUCTION OF TGF-BETA BINDING-PROTEIN MUTANT

##### A. Mutagenesis:

A full-length TGF-beta binding-protein cDNA in pBluescript SK serves  
5 as a template for mutagenesis. Briefly, appropriate primers (see the discussion  
provided above) are utilized to generate the DNA fragment by polymerase chain  
reaction using Vent DNA polymerase (New England Biolabs, Beverly, MA). The  
polymerase chain reaction is run for 23 cycles in buffers provided by the manufacturer  
10 using a 57°C annealing temperature. The product is then exposed to two restriction  
enzymes and after isolation using agarose gel electrophoresis, ligated back into pRBP4-  
503 from which the matching sequence has been removed by enzymatic digestion.  
Integrity of the mutant is verified by DNA sequencing.

##### B. Mammalian Cell Expression and Isolation of Mutant TGF-beta binding-protein:

15 The mutant TGF-beta binding-protein cDNAs are transferred into the  
pcDNA3.1 mammalian expression vector described in EXAMPLE 3. After verifying  
the sequence, the resultant constructs are transfected into COS-1 cells, and secreted  
protein is purified as described in EXAMPLE 3.

20

### EXAMPLE 9

#### ANIMAL MODELS -I

##### GENERATION OF TRANSGENIC MICE OVEREXPRESSING THE *BEER* GENE

The ~200 kilobase (kb) BAC clone 15G5, isolated from the CITB mouse  
25 genomic DNA library (distributed by Research Genetics, Huntsville, AL) was used to  
determine the complete sequence of the mouse *Beer* gene and its 5' and 3' flanking  
regions. A 41 kb Sall fragment, containing the entire gene body, plus ~17 kb of 5'  
flanking and ~20 kb of 3' flanking sequence was sub-cloned into the BamHI site of the  
SuperCosI cosmid vector (Stratagene, La Jolla, CA) and propagated in the *E. coli* strain  
30 DH10B. From this cosmid construct, a 35 kb MluI - AvriI restriction fragment  
(Sequence No. 6), including the entire mouse *Beer* gene, as well as 17 kb and 14 kb of  
5' and 3' flanking sequence, respectively, was then gel purified, using conventional  
means, and used for microinjection of mouse zygotes (DNX Transgenics; US Patent  
No. 4,873,191). Founder animals in which the cloned DNA fragment was integrated  
35 randomly into the genome were obtained at a frequency of 5-30% of live-born pups.  
The presence of the transgene was ascertained by performing Southern blot analysis of

genomic DNA extracted from a small amount of mouse tissue, such as the tip of a tail. DNA was extracted using the following protocol: tissue was digested overnight at 55°C in a lysis buffer containing 200 mM NaCl, 100 mM Tris pH8.5, 5 mM EDTA, 0.2% SDS and 0.5 mg/ml Proteinase K. The following day, the DNA was extracted once with phenol/chloroform (50:50), once with chloroform/isoamylalcohol (24:1) and precipitated with ethanol. Upon resuspension in TE (10mM Tris pH7.5, 1 mM EDTA) 8-10 ug of each DNA sample were digested with a restriction endonuclease, such as EcoRI, subjected to gel electrophoresis and transferred to a charged nylon membrane, such as HyBondN+ (Amersham, Arlington Heights, IL ). The resulting filter was then hybridized with a radioactively labelled fragment of DNA deriving from the mouse *Beer* gene locus, and able to recognize both a fragment from the endogenous gene locus and a fragment of a different size deriving from the transgene. Founder animals were bred to normal non-transgenic mice to generate sufficient numbers of transgenic and non-transgenic progeny in which to determine the effects of *Beer* gene overexpression. For these studies, animals at various ages (for example, 1 day, 3 weeks, 6 weeks, 4 months) are subjected to a number of different assays designed to ascertain gross skeletal formation, bone mineral density, bone mineral content, osteoclast and osteoblast activity, extent of endochondral ossification, cartilage formation, etc. The transcriptional activity from the transgene may be determined by extracting RNA from various tissues, and using an RT-PCR assay which takes advantage of single nucleotide polymorphisms between the mouse strain from which the transgene is derived (129Sv/J) and the strain of mice used for DNA microinjection [(C57BL5/J x SJL/J)F2].

#### ANIMAL MODELS - II

##### 25      DISRUPTION OF THE MOUSE BEER GENE BY HOMOLOGOUS RECOMBINATION

Homologous recombination in embryonic stem (ES) cells can be used to inactivate the endogenous mouse *Beer* gene and subsequently generate animals carrying the loss-of-function mutation. A reporter gene, such as the *E. coli*  $\beta$ -galactosidase gene, was engineered into the targeting vector so that its expression is controlled by the endogenous *Beer* gene's promoter and translational initiation signal. In this way, the spatial and temporal patterns of *Beer* gene expression can be determined in animals carrying a targeted allele.

The targeting vector was constructed by first cloning the drug-selectable phosphoglycerate kinase (PGK) promoter driven *neomycin-resistance* gene (*neo*) cassette from pGT-N29 (New England Biolabs, Beverly, MA) into the cloning vector pSP72 (Promega, Madison, WI). PCR was used to flank the PGK*neo* cassette with

bacteriophage P1 loxP sites, which are recognition sites for the P1 Cre recombinase (Hoess et al., PNAS USA, 79:3398, 1982). This allows subsequent removal of the neo-resistance marker in targeted ES cells or ES cell-derived animals (US Patent 4,959,317). The PCR primers were comprised of the 34 nucleotide (ntd) loxP sequence, 15-25 ntd complementary to the 5' and 3' ends of the PGKneo cassette, as well as restriction enzyme recognition sites (BamHI in the sense primer and EcoRI in the anti-sense primer) for cloning into pSP72. The sequence of the sense primer was 5'-

AATCTGGATCCATAACTTCGTATAGCATACATTATACGAAGTTATCTGCAG  
 10 GATTTCGAGGGCCCCCT-3' (SEQ ID NO:34); sequence of the anti-sense primer was 5'-AATCTGAATTCCACCGGTGTTAATTAAATAACTTCGT  
 ATAATGTATGCTATACGAAGTTATAGATCTAGAG TCAGCTTCTGA-3' (SEQ ID NO:35).

The next step was to clone a 3.6 kb XhoI-HindIII fragment, containing  
 15 the *E. coli*  $\beta$ -galactosidase gene and SV40 polyadenylation signal from pSV $\beta$  (Clontech, Palo Alto, CA) into the pSP72-PGKneo plasmid. The "short arm" of homology from the mouse *Beer* gene locus was generated by amplifying a 2.4 kb fragment from the BAC clone 15G5. The 3' end of the fragment coincided with the translational initiation site of the *Beer* gene, and the anti-sense primer used in the PCR  
 20 also included 30 ntd complementary to the 5' end of the  $\beta$ -galactosidase gene so that its coding region could be fused to the Beer initiation site in-frame. The approach taken for introducing the "short arm" into the pSP72- $\beta$ gal-PGKneo plasmid was to linearize the plasmid at a site upstream of the  $\beta$ -gal gene and then to co-transform this fragment with the "short arm" PCR product and to select for plasmids in which the PCR product  
 25 was integrated by homologous recombination. The sense primer for the "short arm" amplification included 30 ntd complementary to the pSP72 vector to allow for this recombination event. The sequence of the sense primer was 5'-ATTTAGGTGACACT ATAGAACTCGAGCAGCTGAAGCTTAACCACATGGTGGCTCACAACCAT-3' (SEQ ID NO:36) and the sequence of the anti-sense primer was 5'-  
 30 AACGACGGCCAGTGAATCCGTA ATCATGGTCATGCTGCCAGGTGGAGGAGGGCA-3' (SEQ ID NO:37).

The "long arm" from the *Beer* gene locus was generated by amplifying a 6.1 kb fragment from BAC clone 15G5 with primers which also introduce the rare-cutting restriction enzyme sites SgrAI, FseI, AscI and PacI. Specifically, the sequence  
 35 of the sense primer was 5'-ATTACCACCGGTGACACCCGCTTCCTGACAG-3' (SEQ ID NO:38); the sequence of the anti-sense primer was 5'-

ATTACTTAATTAAACATGGCGCGCCAT

ATGGCCGGCCCCCTAATTGCGGCGCATCGTTAATT-3' (SEQ ID NO:39). The resulting PCR product was cloned into the TA vector (Invitrogen, Carlsbad, CA ) as an intermediate step.

5 The mouse *Beer* gene targeting construct also included a second selectable marker, the *herpes simplex virus 1 thymidine kinase* gene (HSVTK) under the control of rous sarcoma virus long terminal repeat element (RSV LTR). Expression of this gene renders mammalian cells sensitive (and inviable) to gancyclovir; it is therefore a convenient way to select against neomycin-resistant cells in which the  
10 construct has integrated by a non-homologous event (US Patent 5,464,764). The RSVLTR-HSVTK cassette was amplified from pPS1337 using primers that allow subsequent cloning into the FseI and Ascl sites of the "long arm"-TA vector plasmid. For this PCR, the sequence of the sense primer was 5'-ATTACGGCCGGCCGCAAAGGAATTCAAGA TCTGA-3' (SEQ ID NO:40); the  
15 sequence of the anti-sense primer was 5'-ATTACGGCGCGCCCCCTC ACAGGCCGCACCCAGCT-3' (SEQ ID NO:41).

The final step in the construction of the targeting vector involved cloning the 8.8 kb SgrAI-Ascl fragment containing the "long arm" and RSVLTR-HSVTK gene into the SgrAI and Ascl sites of the pSP72-"short arm"- $\beta$ gal-PGKneo  
20 plasmid. This targeting vector was linearized by digestion with either Ascl or PacI before electroporation into ES cells.

### EXAMPLE 10

#### 25 ANTISENSE-MEDIATED BEER INACTIVATION

17-nucleotide antisense oligonucleotides are prepared in an overlapping format, in such a way that the 5' end of the first oligonucleotide overlaps the translation initiating AUG of the *Beer* transcript, and the 5' ends of successive oligonucleotides occur in 5 nucleotide increments moving in the 5' direction (up to 50 nucleotides  
30 away), relative to the *Beer* AUG. Corresponding control oligonucleotides are designed and prepared using equivalent base composition but redistributed in sequence to inhibit any significant hybridization to the coding mRNA. Reagent delivery to the test cellular system is conducted through cationic lipid delivery (P.L. Felgner, *Proc. Natl. Acad. Sci. USA* 84:7413, 1987). 2 ug of antisense oligonucleotide is added to 100 ul of  
35 reduced serum media (Opti-MEM I reduced serum media; Life Technologies, Gaithersburg MD) and this is mixed with Lipofectin reagent (6 ul) (Life Technologies,



Gaithersburg MD) in the 100 ul of reduced serum media. These are mixed, allowed to complex for 30 minutes at room temperature and the mixture is added to previously seeded MC3T3E21 or KS483 cells. These cells are cultured and the mRNA recovered. Beer mRNA is monitored using RT-PCR in conjunction with Beer specific primers. In  
5 addition, separate experimental wells are collected and protein levels characterized through western blot methods described in Example 4. The cells are harvested, resuspended in lysis buffer (50 mM Tris pH 7.5, 20 mM NaCl, 1mM EDTA, 1% SDS) and the soluble protein collected. This material is applied to 10-20 % gradient denaturing SDS PAGE. The separated proteins are transferred to nitrocellulose and  
10 the western blot conducted as above using the antibody reagents described. In parallel, the control oligonucleotides are added to identical cultures and experimental operations are repeated. Decrease in Beer mRNA or protein levels are considered significant if the treatment with the antisense oligonucleotide results in a 50% change in either instance compared to the control scrambled oligonucleotide. This methodology enables  
15 selective gene inactivation and subsequent phenotype characterization of the mineralized nodules in the tissue culture model.

## SEQUENCES

Sequence ID No. 1: Human BEER cDNA (complete coding region plus 5' and 3' UTRs)

5  
AGAGCCTGTGCTACTGGAAGGTGGCGTGGCCCTCCTCTGGGTGGTACCATGCAGCTCCCACTGGCCCTGTGTCTCTCTGC  
CTGCTGGTACACACAGCCTTCGGTGTAGTGGAGGGCCAGGGGTGGCAGGCGTTCTAGATGATGCCACGGAAATCATCCC  
CGAGCTGGGAGAGTACCCCGAGCCTCCACCGGAGCTGGAGAAACACAGACCTGAAACGGGGGAGAAACGGAGGGGGC  
CTCCCAACCAACCCCTTTGAGAACCAAGAGCGTGTCCGAGTACAGCTGCCCGGAGCTGCACCTTCAACCGCTACCTGACCGAT  
ORF 10 GGGCGGTGCGCGAGGGCCAGCGGCTCACCAGCTGCTGTGCTCCGGCCAGTGGGSCCGGGCGCGCTCTGCTGCTCAACGC  
CATGGGCGCGGGCAAGTGGTGGCGACCTAGTGGGCGCGACTTCCGCTGCATCCCGAACCGCTACCGCGCGCAAGCGGTGC  
AGGTGCTGTGTCCCGGTGGTGGGCGCGCGCGCGCGCAAGGTGGCGCTGGTGGCGCTGGTGGCGCTGGTGGCGCTGGTGGCG  
CGCTTCCACACCAAGTGGGAGCTCAAGGACTTCCGGAGCCAGGCGCGCTGGCGCGCAAGGCGCGGAGAGCGCGCGCGCGCG  
CGCGCGGAGCGCGCAAGCCACAGCGCGGAGCTGGAGAACCGCTACTAGTGGCGCGCGCGCGCGCGCTCCCGACCGCGCGCG  
15 GCGCGCGCGCGCTGACCGCGCGCGCGCGCGCTTCTGTCTCTGCGCGTGGTTTGATTGTTTATATTTTATTGTATGCGCTGC  
AAGCCAGGGCGAGGGGGCTGAGACCTTCCAGGCGCTGAGGAATCCCGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG  
AGGGGTCCACAGGGCGAGGGGAGGGGATTGAGAGTCAAGACACTGAGCCACGCAGCGCGCGCTCTGGGCGCGCGCTACCT  
TTGCTGGTCCCACTTCCAGGAGGCGAGGAATGGAAGCATTTTCCCGCGCTGGGCTTTTAAAGGAGCGGTGTGGAGTGG  
GAAAGTCCAGGGAGTGGTTAAGAAAGTTGGATAGATTCCCCCTTGCACTCGCTGCCCATCAGAAAGCTGAGGCGTGC  
20 CCAGAGCACAGACTGGGGGCAACTGTAGATGTGTTTCTAGTCTGGCTCTGCCACTAAGTGTGTGTAACCTTGAAC  
TACAGAATTCTCCTTCGGGACCTCAATTTCCACTTTTGAATATGAGGGTGGAGGTGGGAATAGGATCTCGAGGAGACTAT  
TGGCATATGATTCCAAGGACTCCAGTGCCTTTTGAATGGGCGAGGTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG  
CAGTTGCATTGATTCAAGTGCAGGTCACCTTCCAGAAATCAGAGTTGTGATGCTCTCTTCTGACAGCCAAAGATGAAAAA  
CAAAACAGAAAAAAGTAAGAGTCTATTTATGGCTGACATATTTACGGCTGACAAACTCCTGGAGAGAGCTATGCTG  
25 CTTCACAGCCTGGCTTCCCGGATGTTTGGCTACCTCCACCCCTCCATCTCAAAGAAATACATCATCCATTGGGGTAGA  
AAAGGAGAGGGTCCGAGGGTGGTGGGAGGGATAGAAATCAGATCCGCCCCAAGTCCCAAGAGCAGCATCCCTCCCCCG  
ACCCATAGCCATGTTTTAAAGTCACCTTCCGAAGAGAAGTGAAGGTTCAAGGACACTGGCCTTGAGGCGCGGAGGAGC  
AGCCATCACAAGCTCACAGACCAGCACATCCCTTTGAGACACCGCCTTCTGCCACCACTCACGGACACATTTCTGCCT  
AGAAAAAGAGCTTCTTACTGCTCTTACATGTGATGGCATATCTTACACTAAAGAAATATTATTGGGGGAAAAACTAGAGT  
30 GCTGTACATATGCTGAGAACTGCAGAGCATAATAGCTGCCACCCAAAAATCTTTTGAATATCATTTCCAGACACCTC  
TTACTTTCTGTAGTTTTTAATTGTTAAAAAAGGTTTTTAAACAGAGCACATGACATATGAAGCCTGCAGGACT  
GGTCGTTTTTTTGGCAATTCTTCCACGTGGGACTTGTCCACAAGAAATGAAGTAGTGGTTTTTAAAGAGTTAAGTTACAT  
ATTTATTTTCTCACTTAAGTTATTTATGCAAAAGTTTTTCTGTAGAGAAATGACAAATGTTAATATTGCTTTATGAATTAA  
CAGTCTGTTCTTCCAGAGTCCAGAGACATTGTTAATAAAGACAATGAATCATGACCGAAAG  
35

Sequence ID No. 2: Human BEER protein (complete sequence)

MQLFALGLVCLLVHTAFRVVEGGQWQAFKNDATETIFELGEYFEFFFELENNKTMNRAENGGRFHHFFETHDVSEYSC  
 RELHETRYVTGGFGRSAKFVTELVCSSGCGFARLLFNAIGRGMWRPSSGDFRCIFDRYRAQRVQLLCFGGEAFRAHVR  
 5 LVASCKCKRLTRFHNQSELKDFGTEAARPQKGRKPRPRARSAKANQAELENAY

Sequence ID No. 3: Human Beer cDNA containing Sclerosteosis nonsense  
10 mutation

AGAGCCTGTGCTACTGGAAGGTGGGCTSCCCTCCTCTGGCTGGTACCATGCAGCTCCCACTGGGCTGTGTCTGTCTGC  
 CTGCTGGTACACACAGCCTTCCGTGTAGTGGAGGGCTAGGGGTGGCAGGCTTCAAGAATGATGCCACGGAAATCATCC  
 CGAGCTCGGAGAGTACCCCGAGCCTCCACCGAGCTGGAGAACACAGACCATGAACCGGGCGGAGAACGGAGGGCGGC  
 15 CTCCCCACACCCCTTTGAGACCAAGAGCTGTCCGAGTACAGCTGCCGCGAGCTGCACTTCACCCGCTACGTGACCGAT  
 GGGCGGTGCCGAGCGCAAGCCGGTCCCGAGCTGTGTGTCTCCGCGCAGTGCGCCCGGGCGGCTGTGCCCAACGC  
 CATCGGCCCGGGCAAGTGGTGGCGACCTAGTGGGCGGCACTTCCGCTGCATCCCGACCGCTACCGCGGCGAGCGCGTGC  
 AACTGCTGTGTCCCGGTGGTGAAGCGCGCGCGCGCAAGGTGGCGCTGTGGCTCTGTCAAGTGCAGCGCTCACC  
 CGTTCCACAACCACTCGGAGCTCAAGGACTTCGGGACCGAGGCGGCTCGGCGCAGAGGGCGGGAAGCGCGGCGCG  
 20 CGCCCGGAGCGCCAAAGCCAACAGGCGGAGCTGGAGAACGCTACTAGAGCCCGCGCGCGCCCTCCCCACCGCGCGGC  
 GCGCGCGCCCTGAACCGCGCGCCACATTTCTGTCTCTGCGCGTGGTTTGATTGTTTATATTTCATTGTAAATGCCTGC  
 AAGCCAGGGCAGGGGGCTGAGACCTTCAGGCGCTGAGGAATCCCGGGCGCGGCAAGGCCCCCTCAGCCCGCCAGCTG  
 AGGGGTCCACGGGGCAGGGGAGGGATTGAGAGTACAGACACTGAGCCACCGAGCCCGCCTCTGGGCGCGCTACCT  
 TTGCTGGTCCCACTTCAGAGGAGGCAGAAATGGAAGCATTTTACCGCCTGGGTTTAAAGGAGCGGTGTGGGAGTGC  
 25 GAAAGTCCAGGAGTGGTTAAGAAAGTTGGATAAGATTCCCCCTTGACCTCGCTGCCCATCAGAAAGCCTGAGGCGTGC  
 CCAGAGCACAGACTGGGGCAACTGTAGATGTGGTTTCTAGTCTGGCTCTGCCACTAACTTGCTGTGTAACTTGAAC  
 TACACAATTCTCCTTCGGGACCTCAATTTCCACTTTGTAAATGAGGGTGGAGGTGGGAATAGGATCTCGAGGAGACTAT  
 TGGCATATGATTCCAAGGACTCCAGTSCCTTTTGAATGGGCGAGGTGAGAGAGAGAGAGAAAGAGAGAGATGAATG  
 CAGTTGCATTGATTCAAGTCCCAAGGTCACTTCCAGATTGAGAGTTGTGATGCTCTCTTCTGACAGCCAGATGAAGAA  
 30 CAAACAGAAAAAAGTAAGAGTCTATTATGGCTGACATATTACGGCTGACAACTCCTGGAGAAAGCTATGCTG  
 CTCCCCAGCCTGGCTTCCCGGATGTTTGGCTACCTCCACCCCTCCATCTCAAAGAAATAACATCATCCATTGGGGTAGA  
 AAAGGAGAGGGTCCGAGGGTGGTGGGAGGATAGAAATCAGATCCGCCCCCACTTCCCAAAGAGCAGCATCCCTCCCCCG  
 ACCCATAGCCATGTTTTAAAGTCACCTTCGAAGAGAGTGAAGGTTCAAGGACACTGGCCTTGACAGCCGAGGGAGC  
 AGCCATCACAACTCACAGACCAGCAGATCCCTTTTGAAGACCGCCTTCTGCCACCACTCACGGACACATTTCTGCCT  
 35 AGAAACAGCTTCTTACTGCTCTTACATGTGATGGCATATCTTACACTAAAGAATATTATTGGGGGAAAACTACAAGT  
 GCTGTACATATGCTGAGAACTGCAGAGCATAATAGCTGCCACCCAAAAATCTTTTGAAGATCATTTCCAGACAACCTC

TTAATTTCTGTGTAATTTTTTAATTTGTTAAAAAAGAAATTTTAAACAGAGAGCAGCATGACATATGAAAGCCTGCAGGACT  
GGTGGTTTTTTTGGCAATTCTTCACGGTGGGACTTGTTCACAGGAATGAAATAGTAGTGGTTTTTAAAGAACTTAAGTTACAT  
ATTTATTTTCTCACTTAAGTTATTTATGCAAAAGTTTTCTGTAGAGAAATGACAATGTTAATATTCTTTATGAATTAA  
CAGTGTGTTCTTCAGAGTCCAGAGACATTGTTAATAAAGACAATGAATCTGACCCGAAAG

5

Sequence ID No. 4: Truncated Human Beer protein from Sclerosteosis

HQLELALCLVOLLVHTAERVVEG\*

10

Sequence ID No. 5: Human BEER cDNA encoding protein variant (V10I)

15 AGAGCCTGTGCTACTGGAAGGTGGCGTGCCCTCCTCTGGCTGGTACGATGAGAGCTCCCACTGGCCCTGTGTCTCATCTGC  
CTGCTGATACACACAGCCTTCCGTGTAGTGGAGGGGCCAGGGGTGACAGGGCTTCAGAAATGATGACACGGAAATCATCCG  
CGAGCTGGGAGAGTACCCCGAGCCTCCACGGAGCTGGAGAACAACAAGACCATGAACCGGGGCGAGAACGGAGGGGCGGC  
CTCCCAACCAACCCCTTTGAGACCAAGAGCTGTCCAGTACAGCTGCCGCGAGCTGCACCTTACCCCGCTACGTGACCGA  
GGGCCGTGCCGAGCGCCAAGCCGCTCACCGAGCTGGTGTGCTCCGACCACTGCGGCCCGCGCGCTGTGTGCCCAACCG  
20 CATCGGCCGCGGCAAGTGGTGGCGACCTAGTGGGCCCGACTTCCGCTGCATCCCGACCGCTACCGCGCGCAGCGCTGC  
AGCTGCTGTGTCCCGGTGGTGAGGGCGCGCGCGCGCAAGGTGCGCCTGGTGGCCTCGTSCAAGTGCAAGCGCCTCAC  
CGTTCCACAACAGTGGGAGCTCAAGGACTTCGGGACCGAGGCCGCTCGGCCGACAGAGGCCCGGAAGCCGCGGCCCG  
CGCCCGGAGCGCCAAAGCCAACCGAGGCGGAGCTGGAGAACGCTACTAGAGCCCGCCGCGCCCTCCCGACCGGCGGGC  
GCCCGGCCCTGAACCCGCGCCCGACATTTCTGTCTCTGCGCGTGGTTTGATTGTTTATATTTCAATTGTAAATGCCTGC  
25 AACCCAGGGCAGGGGGCTGAGACCTTCCAGGCCCTGAGGAATCCCGGCGCGCGGCAAGGCCCGCCCTCAGCCCGCCAGCTG  
AGGGGTCCACGGGGCAGGGGAGGGAATTGAGAGTACAGAGACTGAGCCACGCAGCCCGCCCTCTGGGGCGCCTACTCT  
TTGCTGGTCCCACTTCAGAGGAGGCAAAATGGAAGCATTTTCACCGCCCTGGGTTTTTAAGGGAGCGGTGTGGGAGTGG  
GAAAGTCCAGGGACTGGTTAAGAAAGTTGGATAAGATTCCCCCTTGACCTCGCTGCCCATCAGAAAGCCTGAGGGGTGC  
CCAGAGCACAAAGACTGGGGGCACTGTAGATGTGGTTTCTAGTCTGGCTCTGCCACTAATTGCTGTGTAACTTGAAC  
30 TACACAATTCTCCTTCGGGACCTCAATTTCCACTTTGTAAATGAGGCTGGAGGTGGGAATAGGATCTCGAGGAGACTAT  
TGGCATATGATTCCAAGGACTCCAGTGCCTTTTGAATGGGCAGAGGTGAGAGAGAGAGAGAAAGAGAGAGAATGAATG  
CAGTTGCATTGATTCAGTGCCAAGGTCACTTCCAGATTTCAGAGTTGTGATGCTCTCTCTGACAGCCAAAGATGAALAA  
CAACAGAAALAAALAAAGTAAAGAGTCTATTTATGGCTGACATATTTACGGCTGACAAACTCTTGAAGAAGCTATGCTG  
CTTCCAGCCTGGCTTCCCGGATGTTTGGCTACCTCCACCCCTCCATCTCAAAGAATAACATCATCCATTGGGGTAGA  
35 AAAGGAGAGGGTCCGAGGGTGGTGGGAGGATAGAAATCACATCCGCCCAACTTCCCAAAGAGCAGCATCCCTCCCCG  
ACCCATAGCCATGTTTTAAAGTCACCTTCCGAAGAGAAGTCAAGGTTCAAGGACACTGGCCTTGACGGCCCGAGGGAGC

ASOCATCAGAACTCAGAGACGACATCCCTTTTGGAGACACCGGCTTTGTCGACACGCTCAAGGACACATTTTGTGCT  
 AGAALACAGCTTCTTACTGCTCTTACATGTGATGGCATATCTTACACTAAAGGATATTATTGGGGGAAAACTACAGT  
 GCTGTACATATGCTGAGAACTGCAGAGCATAATAGCTGCCACCCAAATCTTTTTGAAATCATTTCCAGACACCTC  
 TTAATTTCTGTGAGTTTTTAATTGTTAAAAAAAAGTTTTTAAACAGAGACATGACATATGAAGGCTGCGGAGT  
 5 GGTGTTTTTTTGGCAATTCTTCCAGTGGGACTTGTCCACAGAAATGAAGTAGTGTTTTTTAAAGAGTTAAGTTACAT  
 ATTTATTTTCTCACTTAAGTTATTTATGCAAAAGTTTTTCTGTAGAGAAATGACAATGTTAATATTGCTTTATGATTTA  
 CAGTCTGTTCTTCCAGAGTCCAGAGACATTGTTAATAAGACAAATGAATCATGACCGAAG

10 **Sequence ID No. 6: Human BEER protein variant (V10I)**

MQLFLALCLICLLVHTAFRVVGGWQAFKNDATEIIRELGEYEEFFEELENNKMTMNRANGGEEHHEFFETKDVGEYSC  
 RELHETRYVTDFCRSAKFVTELVCSGQCGEARLLFNAIGRGKWWRESGFDFRCIFDRYRAQRVQLLCFGGEAFRRKRV  
 LVASCKCKRLTRFHNQSELKDFGTEAAREFQNGRKRERERARSAAHQAENAY

15

**Sequence ID No. 7: Human Beer cDNA encoding protein variant (P38R)**

20 AGAGCCTGTGCTACTGGAAGGTGGCGTGCCCTCCTCTGGCTGCTACCATGCAGCTCCCACTGGCCCTGTGTCTGCTCTGC  
 CTGCTGGTACACACAGCCTTCCGTGTAGTGGAGGGCCAGGGTGGCAGGCGTTCAAGAATGATGCCACGGAAATCATCCG  
 CGAGCTCGGAGAGTACCCCGAGCCTCCACCGGAGCTGGAGAACAAACAGACCATGAACGGGGCGGAGAACGGAGGGCGGC  
 CTCCCCACCACCCCTTTGAGACCAAGACGTGTCCGAGTACAGCTGCCGCGAGCTGCACTTCACCCGCTACGTGACCGAT  
 GGGCGGTGCGGACGCGCAAGCCGGTCACCGAGCTGGTGTGCTCCGGCCAGTGGCGCCCGGCGCGCTGCTGCCCAACGC  
 25 CATCGGCCGCGGCAAGTGGTGGCGACCTAGTGGGCGGACTTCGGCTGCATCCCCGACCGCTACCGCGCGCAGCGCGTGC  
 AGCTGCTGTGTCCCGGTGGTGAAGCGCCGCGCGCGCAAGGTGCGCCTGGTGGCCTCGTGCAAGTGCAAGCGCCTCACC  
 CGCTTCCACAACCACTCGGAGCTCAAGGACTTCGGGACCGAGGCGGCTCGGCCGAGAAAGGCGGGAAGCCGCGGCCCG  
 CGCCCGGAGCGCAAGGCCAACAGGCCGAGCTGGAGAACGCCTACTAGAGCCCGCCGCGCCCTCCCCACCGCGGGG  
 GCGCCCGGCCCTGAACCGCGCCCCACATTTCTGTCTCTGCGCGTGGTTTGATTGTTTATATTTCATTGTAAATGCCTGC  
 30 AACCCAGGGGCAAGGGGCTGAGACCTTCCAGGCCCTGAGGAATCCCGGGCGCCGCAAGGCCCCCTCAGCCCGCCAGCTG  
 AGGGGTCCCACGGGGCAGGGGAGGAATTGAGAGTCACAGACACTGAGCCACGCAGCCCCGCTCTGGGGCGGCTACCT  
 TTGCTGGTCCCACTTCAGAGGAGGCAAAATGGAAGCATTTTACCGCCCTGGGGTTTTAAGGGAGCGGTGTGGGAGTGG  
 GAAAGTCCAGGACTGGTTAAGAAAGTTGGATAAGATTCCCCCTTGACCTCGCTGCCCATCAGAAAGCCTGAGGCGTGC  
 CCAGAGCACAAGACTGGGGCAACTGTAGATGTGGTTTCTAGTCTCTGGCTCTGCCACTAATTGCTGTGTAACTTGAAC  
 35 TACACAATTCTCTTGGGACCTCAATTTCCACTTTGTAAATGAGGGTGGAGGTGGGAATAGGATCTCGAGGAGACTAT  
 TGGCATATGATTCCAAGGACTCCAGTGCCTTTTGAATGGGCAGAGGTGAGAGAGAGAGAGAGAAAGAGAGAGAATGAATG

CAGTTGCATTGATTGAGTGCCAAAGTCACTTCCAGGATTCAGAGTTGTGATGTTCTTTTGTGACAGCCAAAGATGAAAA  
 CAAACAGAAAAAAGTAAGATCTATTTATGGCTGACATATTACGGCTGACCACTCTGGAAGAGCTATGCTG  
 CTTCCAGGCTGGCTTCCCCGGATGTTTGGCTAGCTCCACCCCTCCATGTCAAAGAAATACATCATCCATTGGGGTAGA  
 AAGGAGAGGGTCCGAGGGTGGTGGGAGGGATAGAAATCAGATCCGCCCCACTTCCCAAGAGAGCATCCCTCCCCCG  
 5 ACCCATAGCCATGTTTTAAAGTCACTTCCGAAGAGAAGTGAAGGTTCAAGGACACTGGGCTTGCAGGCCGAGGAGT  
 AGCCATCAGAAATCAGAGACCAGGACATCCCTTTGAGACACCGCTTCTGCCCACCACTCAGGGACACATTTTGTGCT  
 AGAAAAAGCTTCTTACTGCTCTTACATGTGATGGCATATCTTACACTAAAGAAATATTATTGGGGGAAATACAGT  
 GCTGTACATATCTGAGAACTGCAGAGCATAATAGCTGCCACCCAAATCTTTTGAATATGATTTCCAGACACCTC  
 TTACTTTCTGTGATGTTTTAATTCTTAAAAAAGTTTTTAAACAGAGCAGATGACATATGAAAGGCTGCAGGACT  
 10 GGTGCTTTTTTGGCAATCTTCCAGGTGGGACTTGTCCACAGAAATGAAAGTAGTGGTTTTTAAGAGTTAAGTTACAT  
 ATTTATTTCTCACTTAAGTTATTTATGCAAAAGTTTTTCTGTAGAGATGACATGTTAATATTGCTTTATGAATTAA  
 CAGTCTGTTCTTCCAGAGTCCAGAGACATTGTTAATAAAGACAATGAATCATGACCGAAAAG

15 **Sequence ID No. 8: Human Beer protein variant (P38R)**

MQLFLALCLVCLLVHTAFRVVEGQSWQLFKNDATETIIRELGEYEEFFEELENNKTMNRAENGGRFFHHFFETHDVSSEYSC  
 RELHETRYVTGDFORSAKEVTELVCSSQCGEARLLFNIGRGKWWRFSGEDFRCIEDRYRAQRVQLLCFSGGEAFRAKVR  
 LVASCKKRLTRFHNQSELKDFGTEAARFQKGRKFRFRARSANQAELEHAY

20

**Sequence ID No. 9: Vervet BEER cDNA (complete coding region)**

25 ATGCAGCTCCCACTGGCCCTGTGTCTTGTCTGCCTGCTGGTACACGCAGCCTTCCGTGTAGTGGAGGGCCAGGGGTGGCA  
 GGCCTTCAAGAATGATGCCACGGAAATCATCCCGAGCTCGGAGAGTACCCGAGCCTCCACCGGAGCTGGAGAACAACA  
 AGACCATGAACCGGGCGGAGAATGGAGGGCGGCCTCCCCACCAACCCCTTTGAGACCAAAGACGTGTCCGAGTACAGCTGC  
 CGAGAGCTGCACTTCACCCGCTACGTGACCGATGGGCCGTGCCGACGCGCAAGCCAGTCACCGAGTTGGTGTGCTCCGG  
 CCAGTGGGGCCCGGCACGCCTGCTGCCAACGCCATCGGCCGCGGCAAGTGGTGGCGCCGAGTGGGCCCGACTTCCGCT  
 30 GCATCCCCGACCGCTACCGCGCGCAGCGTGTGCAGCTGCTGTGTCGGTGGTGGCGCGCGCGCGCGCAAGGTGCGC  
 CTGGTGGCCTCGTGCAAGTGCAAGCGCCTCACCCGCTTCCACAACCAAGTCCGAGCTCAAGGACTTCGGTCCCCGAGGCCG  
 TCGGCCGCAGAAGGGCCGGAAGCCGTGGCCCCGCGCCCGGGGGGCCAAGCCATCAGGCCGAGCTGGAGAAGCCCTACT  
 AG

35

**Sequence ID No. 10: Vervet BEER protein (complete sequence)**

MQLF LALCLVCLLVHAAFRVVEGQSWQAFKNDATETII FELGETFEFFEELENNKTHNRAENGGRFEHHEFETHYDVSEYSC  
RELHETRYVTDSFERSAKFVTELVCSGQCGEARLLENALIRGKWWRFSGEDFRCIIDRYRAQRVQLLCFGGAAPRAKRV  
LVASCKCKRLTRFHINQSELKDFGEFAARFQKGRKPRFRAAGAHANQAELENAY

5

Sequence ID No. 11: Mouse BEER cDNA (coding region only)

10 ATGCAGCCCTCACTAGCCCCGTGCCTCATCTGCCTACTTGTGCAGGCTTCCTTCTGTGCTGTGGAGGGGCGAGGGGTGSCA  
AGCCTTCAGGAATGATGCCACAGAGGTTCATCCAGGGCTTGGAGAGTACCCCGAGCCTCCTCCTGAGAACACACAGACCA  
TGAACCGGGGCGGAGATGGAGGCAGACCTCCCCACCATCCTTATGAGGCCAAGGTGTGTCCGAGTACAGCTGCCGCGAG  
CTGCACTACACCCGCTTCCTGACAGACGGCCCATGCCGAGGCGCAAGCGGTACACGAGTTGCTGTGCTCCGGCCAGT  
CGGCCCCGCGGGCTGCTGCCAAGCCATCGGGCGGTGAAGTGGTGGCGCCGGAAGGACCGGATTTCGGCTGCATCC  
15 CGGATCGCTACCGCGCGAGCGGGTGAGCTGCTGTGCCCGGGGGCGCGCGCGCGCTCCGCGAAGGTGCTGTGCTG  
GCCTGTGCAAGTGCAAGCGCTCACCCGCTTCACACACAGTGGAGCTCAAGGACTTCGGGCGGAGAGCCGCGGGCC  
GCAGAAGGGTCGCAAGCCGCGCGCCCGGGGAGCCAAAGCCAAACAGGCGGAGCTGGAGAACGCTACTAGAG

20 Sequence ID No. 12: Mouse BEER protein (complete sequence)

MQFSLAFCLICLLVHAAPCAVEGQSWQAFRNDATETIIFGLGEYFEFFEEFNQTMNRAENGGRFEHHEFYDAKDVSEYSCRE  
LHYTRFLTDSFERSAKFVTELVCSGQCGEARLLENALIRGKWWRFNGEDFRCIIDRYRAQRVQLLCFGGAAPRAKRVRLV  
ASCKCKRLTRFHINQSELKDFGEFETARFQKGRKPRFAGAGAHANQAELENAY

25

Sequence ID No. 13: Rat BEER cDNA (complete coding region plus 5' UTR)

30 GAGGACCGAGTGCCCTTCCTCCTTCTGGCACCATGCAGCTCTCACTAGCCCCCTTGCCCTTGCTGCTGCTGTACATGCA  
GCCTTCGTTGCTGTGGAGAGCCAGGGGTGGCPAGCCTTCAAGAATGATGCCACAGAAATCATCCGGGACTCAGAGAGTA  
CCCAGAGCCTCCTCAGGAAGTACAGAACACAGACCATGACCGGGCGGAGAGGAGGAGACCCCCCACCATCCTT  
ATGACACCAAGACGTGTCCGAGTACAGCTGCCGCGAGCTGCACTACACCCGCTTCGTGACCGACGGCCCGTGCCGAGT  
GCCAAGCCGCTACCGAGTTGGTGTGCTCGGGCCAGTGGGCCCCGCGGGCTGCTGCCCAACGCCATCGGGCGCGTGAA  
35 GTGGTGGCGCCGGAACGAGCCGACTTCCGCTGCATCCCGATCGCTACCGCGCGCAGCGGTGCAGCTGCTGTGCCCG  
GCGGCGCGCGCCGCGCTCGCGCAAGGTGCGTCTGGTGGCCTCGTGCAAGTGCAAGCGCCTCACCCGCTTCACACACAG

HQLSLAFCLACLLVHAFVAVESQGWQAFNHIDATETIIFGLREYFEFFOELENNHQTMRRAENGGRFFHHHEYDTHDUSEYV  
 RELHYTKREYTDGECKSAKEVTELVCSGGQGFARLLLENAIGRVKWWRENGEDERCIEFDTATRAQRVQLLCFGGAAFAEPTTA  
 LVA.SCHOKALTRFHNQSELKDGEFETARFQGRKRFPRAGAKAHQAELENAV

15 A G A T T G A T G C C A C A G A A T C A T C C C G A G C T G G G C G A G T A C C C G A G C C T C T G C G A G A G C T G A A C A A C A A G A C C A T G A A C  
C G G G C G G A G A A C G G A G G G A G A C C T C C C A C C A C C C C T T T G A G A C C A A A G A C G C C T C G A G T A C A G C T G C C G G G A G C T G C A  
C T T C A C C C G C T A C G T G A C C G A T G G C C G T G C C G C A G C G C C A A G C C G G T C A C C G A G C T G G T G T G C T C G G G C C A G T G C G G C C  
C G G C G C C C T G C T G C C C A A C G C C A T C G G C C G G G C A A G T G G T G G C G C C C A A G C G G G C C G A C T T C C G C T G C A T C C C C G A C  
C C T A C C C G C G C A G C G G T G C A G C T G T T G T G T C C T G G C G C G C G G C G C C G C G C G C G C A A G G T G C C C C T G G T G C C C T C  
20 T T G C A A G T G C A A G C G C C T C A C T C G T T C C A C A A C C A G T C C G A G C T C A A G G A C T T C G G G C C C G A G G C C G C G C G C G C A A A  
C G G G C C G A A G C T G C G G C C C C G C G C C G G G G C A C C A A A G C C A G C C G G G C G A

NDATEIIFELGEYFEELFELNNKTMHRAENGGRFEHFFETKDASEYSCRELHFTRYVTDGFCRSAPVTELVCQGQGF  
AALLFNAIGRGKWWRESGFDFRCIFDRYRAQRVQLLCFGGAERARKVRLVASCKCKRLTRFHNQSELKDFGEAAAEFQT  
GRKLRFRARGTKASRA

CGCGTTTGGTGAGCAGCAATATTGCGCTTCGATGAGCCTTGCGTTGAGATTGATACCTCTGCTGCACAAAAGGCATTG



3ACCGAGCTGACCCAGCGCATTCGTGACACCGTCTCCTTGGAACTTATTGGCAATGGAGTGTCAATTCATCGAGGACNGCC  
TGATCGCAATGGTGTCTATCCACCGCAGCGGCATCGAAAGACCTCAGCGCGTGAACCAATATCTACACATCAGCCTTGST  
ATCCTGCGTGATGAGCCAGCGCAGAACAGGTAAACGTCAGTGGCGATAAATTCAGAGTTAAGCTGTTGATACCA  
CATTGAAGCGTTGATCGAAACCGCGCTGAAAGACGCTGCTGAATGTGGGCGCTGGATGTCAAAAGCAGATGGCAGCAG  
5 ACAAGAAAGCGATGGATGAAGTGGCTTCCTATGTCCGACCGGCATCATGATGGAATGTTTCCCGGTGGTGTATCTGG  
CAGCACTGCCGTGATAGTATGCAATTGATAATTATTATCATTTGCGGCTTCTTTCCGGCGATCCGCTTGTACGGGGC  
GGCGACCTCGCGGGTTTTCGCTATTTATGAAGATTTTCGGSTTTAAGGCGTTTCGCTTCTTCTGTCATAAGTTAATGT  
TTTTATTTAAATACCTCTGAAAGAAAGGAAAGCAGAGTGTGTAAGCGAGCTTTTGGCGCTGTGCTTTCTCTTC  
TCTGTTTTGTCCGTGGAATGAACATGGAGTCAACAAAGCAGAGCTTATCGATGATAAGCGGTCAACATGAGAT  
10 TCGCGCGCGCATATAGACTCACTATAGGGATCGAGCGCTACTCCCGCGCATGAAGCGGAGGAGCTGGACTCCGCTG  
CCGAGAGCGCCCCCAACCCCCAAGTGCCTGACCTCAGCGCTCTAGCGCTCTGGCTTGGGCTTGGGCGGGTCAAGGC  
TACCAAGTTCTCTTAACAGGTGGCTGGCTGTCTCTTGGCGCGCGCTCATGTGACAGCTGCTAGTTCTGCAGTGGGTG  
ACCGTGGATGTCTGCTTCTGCTTGGCATGGCAACGGGATGAGCTTACATCTGGGTGTGGAGCTTTCTCTGCTGTCA  
GGAAATGCAATACCTTAAATACCTAGAGAGGAGTAGCTGAGCGCAAGCTTTCTGGCTTCTCCAGATAAAGTTTG  
15 ACTTAGATGGAAAGAAACAAATGATAAGACCGGAGCATCTGAAATTTCTCTAATTGCACCACTAGGAAATGTGT  
TATTATTAGCTCGTATGTGTCTTATTTTAAAGAAAGCTTTAGTCATGTTATTAATAAGATTTCTCAGCAGTGGGA  
GAGAACCAATATTAAACCAAGATAAAGTTGGCATGATCCACATTCAGGAGAGATCCAGCTTGGGTTTTCTGAATGTG  
AAGACCGCTTTATTAAGTCTAAGCTCTGTTTTTGCACACTAGGAGCGATGGCGGGATGGCTGAGGGGCTGTAAAG  
ATCTTTCAATGTCTTACATGTGTGTTTCTGTCTGACCTAGGACCTGCTGCTAGCTGCAGCAGAGCCAGAGGGGT  
20 TCACATGATTAGTCTCAGACACTTGGGGGCGAGTTGCATGTACTGCATCGCTTATTCCATACGGAGCACCTACTATGTG  
TCAAGACCATATGGTGTCTACTCTCAGAACGGTGGTGGTCAATCATGCTGCTTGTGACGGTTGGATTGGTGGTAGA  
GAGCTGAGATATATGGACGCACTCTTACGATTTCTGTCAACGTGGCTGTGCTTCTGCTCCTGAGCAAGTGGCTAAACA  
GACTCACAGGGTCAGCCTCCAGCTCAGTCGCTGCATAGTCTTAGGGAACCTCTCCAGTCTCCCTACCTCAACTATCCA  
AGAAGCCAGGGGGCTTGGCGGTCTCAGGAGCCTGCTTGTGGGGGACAGGTTGTTGAGTTTTATCTGCAGTAGGTTGCT  
25 AGGCATAGTGTGAGGACTGATGGCTGCCTTGGAGAACACATCCTTTGCCCTCTATGCAATCTGACCTTGACATGGGGC  
GCTGCTCAGCTGGGAGGATCAACTGCATACCTAAGCCAGCCTAAAGCTTCTTGTCCACCTGAACTCCTGGACCAAG  
GGGCTTCCGGCACATCCTCTCAGGCCAGTGAGGAGTCTGTGTGAGCTGCACTTTCCAATCTCAGGGCGTGAGAGGCAGA  
GGGAGTGGGGGACAGCCTTGACGCTCTTCTCTCCATCTGACAGCGCTCTGGCTCAGCAGCCATATGAGCACAGGC  
ACATCCCCACCCACCCACCTTCTCTGTCTGCAAGATTTAGGCTCTGTTACGGGGGGGGGGGGGGGGGAGTCC  
30 TATCCTCTCTTAGGTAGACAGGACTCTGCAGGAGACACTGCTTTGTAAAGTACTGCAGTTTAAATTTGGATGTTGTGAGG  
GGAAAGCGAAGGGCTCTTTGACCATTAGTCAAGGTACCTTCTAACTCCCATCGTATTGGGGGGCTACTCTAGTGCTAG  
ACATTGCAGAGAGCCTCAGAACTGTAGTTACCAGTGTGGTAGGATTGATCCTTCAGGGAGCCTGACATGTGACAGTTCCA  
TTCTTCACCCAGTCACCGAACATTTATTAGTACCTACCCCGTAACAGGCACCGTAGCAGGTACTGAGGGACGGACCACT  
CAAGAACTGACAGACCGAAGCCTTGGAAATATAACACCAAGCATCAGGCTCTGCCAACAGAACTCTTTAACTCA  
35 GGCCCTTTAACTCAGGACCCCCACCCCAAGCAGTGGCACTGCTATCCACATTTACAGAGAGGAAAACTA  
GGCACAGGACGATATAAGTGGCTTGTCTAAGCTTGTCTGCATGGTAAATGGCAGGGCTGGATTGAGACCCAGACATTCCA

ACTGTAGGGTCTATTTTCTTTTCTCGTTGTTGGAATCTGGGTCTTACTGGGTAAATCAGGGCTAGGCTGACACTCAT  
ATCCTTCTCCCATGGCTTACGAGTGTAGGATTCAGGTGTGTGCTACCATGTCTTACTCCCTGTAGCTTGTCTATAGCA  
TCTTCACACATAGGAATTGTGATAGCAGCACACACACCGGAGGAGCTGGGAAATCCACAGAGGGCTCCGACGAGTG  
ACAGGGGAATGCCTACACAGAGGTGGGGAAGGGAAGCAGAGGGGAACAGCATGGGGCTGGGACCACAAGTCTATTTGGGG  
5 AAGCTGCCGGTACCGTATATGGCTGGGGTGGGGGAGAGGTGATGAGATGAGGCAGGAGAGGCCACAGCAGGACAGGGG  
TACGGGCTCCTTATTGCCAGAGGCTCGGATCTTCCTCCTCTTCTCCTCCTCCGGGGGTGCTGTTTCTATTTTCCACCACTG  
CCTCCCATCCAGGTCTGTGGCTCAGGACATCACCCAGCTGCAGAACTGGGCATCAGCCAGCTCTGAATGCTGCGGAGG  
GCAGGTCTTTCATGCAGGTCAACACCCAGTGTAGCTTCTACGAGGATTTCTGGCATCACTACTTGGGCATCAAGGCGCAT  
GATACGAGGAGTTCAACCTCAGTCTTACTTTGAAAGGGCCAGAGATTTCTTGAACAGGGGCTGGCCCATAAATAGG  
10 TAGGAAGCTACATTCGGGCACCCATGGAGCGTAAGCCCTCTGGGACCTGCTTCTCTCAAGAGGGGCCACTTGAAGAA  
GGTTCCAGAAAGATCCCAATATATGCCACCACTAGGGATTAAGTGTCTACATGTGAGCCGATGGGGGCCACTGCATAT  
ACTGTGTCCATAGACATGACAAATGGATAATAATATTTAGACAGAGAGCAGGAGTTAGGTAGCTGTGCTCTTCTCCTT  
TATTGAGTGTGCCATTTTTTTTATTCATGTATGTGTATACATGTGTGTGCACACATGCCATAGGTTGATAGTGAACCC  
GTCTTCAATCTTCCCAACCCCACTTATTTTTGAGGCAGGGTCTCTTCTCTGATCTCTGGGGCTCATTTGGTTTATCTAG  
15 GCTGTGGCCAGTGAGCTCTGGAGTTCTGCTTTTCTCTACCTCCCTAGCCCTGGGACTGCAGGGGCATGTGCTGGGCCAG  
GCTTTTATGTGCGTTGGGGATCTGAACTTAGGTCCCTAGGCCTGAGGACGTTAAGACTCTGCCACATCCCAAGCTGT  
TTGAGCAAGTGAACCATTCGCCAGAAATCCCCAGTGGGGCTTCTCACTCTTTTATTGGCTAGGCATTCATGAGTGGTC  
ACCTCGCCAGAGGAATGAGTGGCCACGACTGGCTCAGGGTCAGCAGCTAGAGTACTGGGTAAAGTCTTCTCGCCGCTC  
GCTCCCTGCAGCCGAGACAGAAAGTAGGACTGAATGAGAGCTGGCTAGTGTGAGACAGGACAGAAGGCTGAGAGGGTC  
20 ACAGGGCAGATGTGAGCAGAGCAGACAGGTTCTCCCTCTGTGGGGAGGGGTGGCCACTGCAGGTGTAATTGGCCTTCT  
TTGTGCTCCATAGAGGCTTCTGGGTACACAGCAGCTTCCCTGTCTGGTGATTCCCAAGAGAACTCCCTACCACTGGA  
CTTACAGAAAGTTCTATTGACTGGTGTACGGTTCAACAGCTTTGGCTCTTGGTGGACGGTGCATACTGCTGTATCAGCTC  
AAGAGCTCATTCGAATGAACACACACACACACACACACACACACACACACACAGCTAATTTTATGATGCCTTAACTA  
GTCAGTGACTGGGCATTTCTGAACATCCCTGAAGTTAGCACACATTTCCCTCTGGTGTCTCTGGCTTAACACCTTCTAA  
25 ATCTATATTTTATCTTTGCTGCCCTGTTACCTTCTGAGAAGCCCCCTAGGGCCACTTCCCTTCGCACCTACATTGCTGGAT  
GGTTTCTCTCCTGCAGCTCTTAAATCTGATCCCTCTGCCTCTGAGCCATGGGAACAGCCCAATAACTGAGTTAGACATAA  
AAGCTCTCTAGCCAAACTTCACTAAATTTAGACAATAATCTTACTGGTTGTGGAATCCTTAAGATTCTTCATGACC  
TCCTTCACATGGCAGAGTATGAAGCTTTATTACAATGTTTATTGATCAAACTAACTCATAAAAGCCAGTTGTCTTTC  
ACCTGCTCAAGGAAGGAACAAAATTCATCCTTAACTGATCTGTGCACCTTGACAAATCCATACGAATATCTTAAGAGTAC  
30 TAAGATTTTGGTTGTGAGAGTCACATGTTACAGAATGTACAGCTTTGACAAGGTGCATCCTTGGGATGCCGAAGTGACCT  
GCTGTTCCAGCCCCCTACCTTCTGAGGCTGTTTTGGAAGCAATGCTCTGGAAGCAACTTTAGGAGGTAGGATGCTGGAAC  
AGCGGGTCACTTCAGCATCCCGATGACGAATCCCGTCAAAGCTGTACATTCTGTAACAGACTGGGAAAGCTGCAGACTTT  
AAGGCCAGGGCCCTATGGTCCCTCTTAACTCCCTGTACACCCACCCGAGCCCTTCTCCTCCAGCCGTTCTGTGCTTCTC  
ACTCTGGATAGATGGAGAACACGGCCTTGTAGTTAAAGGAGTGAGGCTTCAACCTTCTCACATGGCAGTGGTTGGTCAT  
35 CCTCATTCAGGGAAGCTCTGGGGCATTCTGCCTTTACTTCTCTTTTGGACTACAGGGAATATATGCTGACTTGTTTTGA  
CCTTGTGTATGGGAGACTGGATCTTTGGTCTGGAATGTTTCTGCTAGTTTTTCCCCATCCTTTGGCAAACCTATCTA

TATCTTACCACTAGGCATAGTGGGCGCTCTCTCTGAGGCTTGGCTTCAGGCTGGTTCTGGGGAATATCTTCTGCTTCTCT  
CCCCAGCATATGCTGCTCAGAGTGTTCCTCTGGGCTGCTTGTGTAACAAAGCGGGGATTCATCCGAGAGCTCGGGTGGC  
TTGTGGGTACACTGCTAAGATAAATGGATACTGGGCTCTCTCTGACCACTTGCAGAGCTCTGGTCTCTTGTGGGTACAG  
TGCTAAGATAAATGGATACTGGGCTCTCTCTATCCACTTTCAGGAGCTCTAGGGAGCAGGAATCCATTACTGAGAAACCC  
5 AGGGGCTAGGAGCAGGGAGGTAGCTGGGCGAGCTGAATGCTTGGGAGCTAACCAATGAATACAGAGTTTGGATCTCTAG  
AATACTCTTAAATCTGGGTGGGAGAGTGGGCTGCTGTAATCCAGAGCTCGGGAGGCGGAGACAGGGAATCATCAGA  
GCAACTGGCTAACAGAAATAGCAAAACACTGAGCTCTGGGCTCTGTGAGAGATCTGCTTAAACATATAAGAGAGAGAA  
TAAAGCATTAAGAGAGACAGTAGATGCCAATTTTAAAGCCCGACATGCACATGGACAAGTGTGGTCTTGAACACACATAT  
GCACTCATGTGAACAGGAGCATGCACACTCGGGCTTATCAACACATATTTTGAAGAGAGAGTGAAGAGAGAGAGTGCAC  
10 ATTAGAGTTTACAGGAAAGTGTGAGTGAGCACACCCCTGACACAGACATGTGTGCCAGGGAGTAGGAAAGGAGCTGGG  
TTTGTGTATAAGAGGGAGCCATCATGTCTTCTAAGGAGGGCGTGTGAAGGAGGCGTTGTGTGGGCTGGGACTGGAGCAT  
GGTGTGAAGTGAAGCATGTCTCTGTGGGAAACAGGAGGGTGGGCGAGCTGCAGAGGCTCCCACTGTCCAGCGGATCAAT  
AAAAGCCCGCTGTGAGAACTTTAGGTAATAGCCAGAGAGAGAAAGTAGGAAGTGGGGGAGCTCCCATCTCTGATGTAG  
GAGGATCTGGGCAAGTAGAGGTGGCTTTGAGGTAGAAAGGGGCTGCAGAGGAGATGCTCTTAATTTCTGGGTGAGGAGTT  
15 TCTTTCCAAATATGCTGTGAGGAGGTGTAGGTGGTGGCCATTCACTCACTCAGCAGAGGGATGATGATGCCCGGTGGG  
TGCTGGAAATGGCCAGCATCAACCCCTGGCTCTGGAAGAACTCCATCTTTACAGAGGAGAGTGGATCTGTGTATGGCCAG  
CGGGGTACAGGTGCTTGGGGCCCTGGGGGAGCTCCTAGCACTGGGTGATGTTTATCGAGTGCTCTTGTGTGCCAGGCAC  
TGGCCTGGGGCTTTGTTCTCTCTCTGTTTGTCTTCTTTTGGAGACAGACTCTTGTGTATGATCCGTGTCAATCTTGG  
AATCTCACTGCATAGCCAGGCTGCGGAGAGAGGGGAGGGCAATAGGCCCTGTAAAGCAAGCCACACTTCAGAGAGCTAGAC  
20 TCCACCTGCGAATGATGACAGGTGAGAGCTGAGTTCGGGAGATTTTTTTTCCAGCTGCCAGGTGGAGTGTGGAGTGGG  
AGCTAGCGGCAAGGAGTAGAGGGCGAGCTCCCTGTGCAGGAGAAATGCAAGCAAGAGATGGCAAGCCAGTGAGTTAAGCAT  
TCTGTGTGGGAGCAGGTGGATGAAGAGAGAGGCTGGGCTTTGCGCTCTGGGGGGGGGTGAGGGGTGGGATGAGGTGA  
GAGGAGGGGAGCTCCCTGCAGTGTGATGAGATTTTTCTGACAGTGACCTTTGGCCTCTCCCTCCCCACTTCCCTTCTT  
TCCTTTCTTCCACCATTTGCTTTCTTGTCTTGGAGAAATCTGAGTTTCCACTTCACTGGTGATGCAGACGGAAACAGA  
25 AGCCGT  
GTATGTGTGTGAGTGGGAATGGCTCATAGTCTGCAGGAAAGTGGGCAAGGAATAAGCTGTAGGCTGAGGCAAGTGTGG  
GATGCAGGGAGAGAGGAGAGGAGGATACCAGAGAAGGAATTAAGGGAGCTACAAGAGGGCATTTGTTGGGCTGTGTGTG  
TGT  
TGT  
30 AACTGGAGTTGGAGGAGGTTGTGAGTCCCTGACATGTTTGTGGGAACTGAACCCCGGTCTATGCAAGAGCAGGAAGT  
GCAGTTATCTGCTGAGCCATCTCTCCAGTCTGAAATCCATTCTCTTAAATACACGTGGCAGAGACATGATGGGATTTA  
CGTATGGATTTAATGTGGCGGTCAATTAAGTTCCGGCAGAGCAAGCACCTGTAAAGCCATCACCACAACCGCAACAGTGA  
ATGTGACCATCACCCCATGTTCTTCTATGTCCCTGTCCCTCCATCTCTCAAGCACCTCTTGTCTGCGCTCTG  
TCGCTGGAGAACAGTGTGCATCTGCACACTCTTATGTCAAGTGAAGTGCACAGGCTGCACCCCTTCTGGTCTGAGTATT  
35 TGGGTTCTGACTCTGCTATCACACACTACTGTACTGCATTCTCTCGCTCTCTTTTTTAAACATATTTTATTTGTTTGT  
GTGTATGCACATGTGCCACATGTGTACAGATACTATGGAGGCCAGAGAGGCCATGGCCGTCCCTGGAGCTGGAGTTACA

GGCAGCGGTATGAGCTGCTGATGTTGGGTGCTGGGAACCAAGCTTAAATCTAAAGACAGACCTTTAACTCTGAGGAGG  
TCTGAGTACCCCTTCTTCATTTCTCCGOUTGGGTTCCTATGATGACACATGTAGCTAGAAATATCTTGGTTATCTAATTA  
TGTACATTGTTTTTGTCTAAGAGAGAGTAAATGCTCTATAGCCTGAGCTGGCCTCAACCTTGGCATCCTCTCCCTCAGGCT  
TCCTCCTCCTGAGTGTCTAGGATGACAGGGGAGTGGTAACCTTACATGCTTTCATGTTTTGTTCAAGAGCTGAAGGATAACAT  
5 TCATACAGAGAGGCTCTGGGTCAACAGTGTGCACTTCACTGAAATGGCAGCAACCGGTGATCAAGAAACAAAGCTCAGGG  
CTGGAGAGATGGCACTGACTGCTCTCCAGAGGTCCGAGTTTCATTCGACAGCAACCATGCTGGCTCAGAGGCTCT  
TACAGAGATCTGAGGCCCTCTTCTGGTGTGTCTGAAGACAGCTACAGTGTACTCAGATAAATAAATAAATCTTTAAAA  
ACACACACACACACATTAACACGCCAGGAAGGCCACTCCATGTTCCCTCCGAGGTCTCTGCCCTACAGTACTCCAGGTT  
ACCACTGTTCAAGGCTTCTAACAACCTGGTTTTACTTGGGCTCTTTTTCTGCTCTGTGGAGCCACACATTTGTGTGCTCAT  
10 ACAGCTTCTTTCTAGTAAGTTGCATATTACTCTGGGTTTTCTACATGATTTATTTATTTAGTGTGTGTGGCTGTGGG  
CCATGCTATGGCAGAGTGTGTGGGGATGTCAAGATATTGTGAACAGGGGACAGTTCTTTTCTCAATCATGTGGGTTCCAG  
AGGTTGAAGTCAAGTCAATCATGTGTGGCAGCAATGCCCTTACCCACTCAGAGCATCTCCATATTCTTTTTTTTTCCCTG  
AGGTGGGGGCTTGTTCATAGCCCAACTGGCTTTGCACTTGCAGTTCAAAAGTCACTCCCTGTCTCCACCTCTTAGAGTA  
TTGGAAATTACGATGTGTACTACCACACCTGACTGGATCATTAATTCTTTGATGGGGGCGGGGAGCGGCACATGCTGCAG  
15 TGAAGGATGACTGGACTGGACATGAGCGTGGAAAGCCAGAGAACAGCTTCAGTCTAATGCTCTCCCAACTGAGCTATTTC  
GTTTTGCCAGAGAACAACTTACAGAAAGTTCTCAGTGGCATGTGGATTGGGGTTGGAGTCAACTCATCAGCTTGACAT  
TGGCTCTCTACCCACTGAGCCTTCTCACTACTCTCTACCTAGATCATTAATTCTTTTTTAAAGAGCTTATTAGGGGG  
TGGAGAGATGGCTCAGCGTTAAGAGCACCGAATGCCCTTCAGAGGTCCTGAGTTCAATTCGACGATGCCATTGCTGG  
GCAGTAAAGGGGCGCAGGTGTTCAACGTGAGTAGCTGTTGCCAGTTTTCCGCGGTGGAGAACCTCTTGACACCCCTGCTGT  
20 CCTGGTCATTCTGGGTGGGTGCATGGTGATATGCTTGTGTATGGAAGACTTTGACTGTTACAGTGAAGTTGGGCTTCCA  
CAGTTACACAGTCTCCCTGTTTTCTGCAAGGCCGGGTGCTTGTCCATTGCCGCGAGGGCTACAGCCGCTCCCCAACGGCTA  
GTTATCGCCTACCTCATGATGCGGCAGAAGATGGACGTCAAGTCTGCTCTGAGTACTGTGAGGCAGAATCGTGAGATCGG  
CCCCAACGATGGCTTCTGGCCCAACTCTGCCAGCTCAATGACAGACTAGCCAAAGGAGGGCAAGGTGAAGCTCTAGGGTG  
CCCACAGCCTCTTTTGAGAGGCTCTGACTGGGAGGGCCCTGGCAGCCATGTTTAGGAACACAGTATACCCACTCCCTGC  
25 ACCACCAGACAGTGCCACATCTGTCCCACTCTGGTCCCTCGGGGGCCACTCCACCTTAGGGAGCACATGAAGAAGCTC  
CCTAAGAAGTTCTGCTCCTTAGCCATCCTTTCTGTAAATTTATGTCTCTCCCTGAGGTGAGGTTCAAGTTTATGTCCCTG  
TCTGTGGCATAGATACATCTCAGTGACCCAGGGTGGGAGGGCTATCAGGCTGCATGGCCCCGGGACAGGGGCACTCTTCAT  
GACCCCTCCCCACCTGGGTCTTCTCTGTGTGGTCCAGAACCACGAGCCTGGTAAGGAAGTATGCAAAACACAGGGCCCTG  
ACCTCCCCATGTCTGTTCCCTGGTCCCTCACAGCCCGACAGCCCTGCTGAGGCAGACGAATGACATTAGTTCTGAAGCAG  
30 AGTGAGATAGATTAGTGACTAGATTTCCAAAAGAGGAGGAGGAGGCTGCATTTTAAATTTATTTCCCTAGAAATTA  
AGATACTACATAGGGGCCCTTGGGTAAAGCAATCCATTTTTCCAGAGGCTATCTTGATTCTTTGGAATGTTTAAAGTGT  
GCCTTGCCAGAGAGCTTACATCTATATCTGCTGCTTCAAGCCTTCCCTGAGGATGGCTCTGTTCTCTTGTGTTAGTA  
AGAGCGATGCCTTGGGCAGGGTTTTCCCCCTTTTCAGAATACAGGGTGTAAAGTCCAGCCTATTACAAACAAACAAACAA  
CAAACAAACAAAGGACCTCCATTTGGAGAATTGCAAGGATTTTATCCTGAATTATAGTGTGGTGAGTTCAGTCATCAC  
35 GCCAAGTGCTTGCCATCCTGGTTGCTATTCTAAGAATAATTAGGAGGAGGAGCCTAGCCAATTGCAGCTCATGTCCGTGG  
GTGTGTGCACGGGTGCATATGTTGGAAGGGGTGCTGTCCCTTGGGGACAGAAGGAAATGAAAGGCCCTCTGCTCAC





ATGAGGGAATGATTTTTTTGTAAAGATGAAATTTTGTGTTGGGTCGAAAGAGGCTGGGCAAGGAGGAATGCTTTT  
GCACACCAAGGCTATAAGTCACCATGAGTTCCCTGGCTAAGATCAGATGTAAATGGAGGCGAGGCTCCCTCTGCTGGTGG  
TTGCTCTCTCCACTGGTTTTGAAGAGAAATTCAGAGAGATCTCCTTGGTCAGAAATTGTAGGTGCTGAGCAATGTGGAGC  
TGGGGTCATGGGATTCCCTTTAAGGCATCCTTCCCAAGGCTGGGTCTACTTCAATAGTAGGGTGGCTGGCAGGCAAGC  
5 GTGAGACCCCTAGGTAGAGTCCCCAGAACTGCCCCCAACCCCCCAAGAGCATCCTTCTGCTCTGGGTGGGTGGG  
GAGCAACACCTTTAACTAGACCATTAGCTGGCAGGGTAAACAATGACCTTGGCTAGAGGAATTTGGTCAAGCTGGGT  
TCCGCTTTCTGTAGAGGCCCACTTGTTCCTTTGTTAAGCTGGCCCAAGTTTGTTTTGAAGATGCTGAGGGGCCAG  
GGAGCCAGACAATTAAGAGCAAGCTCATTTTGATATGTAAAGCAGAGGCTGACTGCTTCCCTGGGAGGTACTGG  
GAGAGCTGCTGTGCTCCCTGCTCAGCAAGCCCCCAGCCCAAGCACTCCTGGGTCACTGGGAGGTGGCAGCAG  
10 CAATTTGGAATTTACTGAGCTTGAGAGTCTTGGGAGGGGTGAGCTAAGCAACCCCTTCTCCAGCCCCCCCCACCC  
ACCCCTGTGAGGAGGAGGTGAGGAAACATGGGACCAAGCCCTGCTCCAGCCCTCCTTATTGCTGGCATGAGGAGAGG  
GGGCTTTTAAAGGCAACCTATCTAGGCTGGACACTGGAGCCTGTCTACCAAGTCCCTCCTCCAGCTGGCAGCATGC  
AGCCCTCACTAGCCCCGTGCTCATCTGCTACTTGTGCAAGGTGCTTCTGTGCTGTGGAAGGCCAGGGGTGGCAAGCC  
TTGAGGAATGATGCCACAGAGGTCACTCCAGGGCTTGGAGAGTACCCAGAGCCTCCTCTGAGAACACAGAGCATGAA  
15 CCGGGCGGAGAATGGAGGCAGACCTCCCCACCATCCTATGAGCCCAAGGTACGGGATGAGAGAGCACATTAGTGGGG  
GGGGGTCTGGGAGGTGACTGGGTGGTTTTAGCATCTTCTTCAGAGGTTTGTGTGGGTGGCTAGCCTCTGCTACATCA  
GGGACAGGACACATTTGCTGGAAGAATACTAGCACAGCATTAGAACCCTGGAGGGCAGCATTGGGGGGCTGGTAGAGAGC  
ACCCAGGCGAGGTGGAGGTGAGGTGAGCCGAGGTGCAATTAACAGGGCATGGGCTTGTATGATGGTCCAGAGATC  
TCCTCCTAGGATGAGGACACAGGTGAGATCTAGCTGCTGACCAGTGGGGAAGTGAATGCTGAGGCTGGATGCCAGATG  
20 CCATCCATGGCTGTACTATATCCACATGACCACCACATGAGGTAAGAGAGGCCCCAGCTTGAAGATGGAGAAACCGAGA  
GGCTCCTGAGATAAAGTCACTGGGAGTAAGAAGAGCTGAGACTGGAGGCTGGTTTGATCCAGATGCAAGGCAACCCCTAG  
ATTGGGTTTGGGTGGGACCTGAAGCCAGGAGGAATCCCTTTAGTTCCCTCTGCCAGGGTCTGCTCAATGAGCCAGAG  
GGGTAGCATTAAAGAGACAGGGTTTGTAGGTGGCATGTGACATGAGGGGCAGCTGAGTGAATGTCCCTGTATGAGCA  
CAGGTGGCACCCTTGCCTGAGCTTGACCCCTGACCCCAAGCTTTGCTCATTCCTGAGGACAGCAGAACTGTGGAGGC  
25 AGAGCCAGCACAGAGAGATGCTGGGGTGGGGTGGGGTATCACGACGGAACTAGCAGCAATGAATGGGTGGGTGG  
CAGCTGGAGGGACACTCCAGAGAAATGACCTTGTGCTGCTGACCATTTGTGTGGGAGGAGGCTCATTTCCAGCTTGCCAC  
CACATGCTGTCCCTCCTGTCTCCTAGCCAGTAAGGGATGTGGAGGAAGGGCCACCCCAAGGAGCATGCAATGCAGTCA  
CGTTTTTGCAGAGGAAGTCTTGACCTAAGGGCACTATTCTTGGAAAGCCCCAAACTAGTCTTCCCTGGGCAACAGG  
CCTCCCCACATACCACCTCTGCAGGGGTGAGTAATTAAGCCAGCCACAGAGGGTGGCAAGGCTACACCTCCCCCT  
30 GTTGTGCCCCCCCCCCCCCGTGAAGGTGCATCCTGGCTCTGCCCTCTGGCTTGGTACTGGGATTTTTTTTTTCTTT  
TTATGTCAATTTGATCCTGACACCATGGAATTTTGGAGGTAGACAGGACCCACACATGGATTAGTTAAAGCCTCCCAT  
CCATCTAAGCTCATGGTAGAGATAGAGCATGTCCAAGAGAGGAGGGCAGGCATCAGACCTAGAAGATATGGCTGGGCAT  
CCAACCCAATCTCCTTCCCCGGAGAACAGACTCTAAGTCAGATCCAGCCACCTTGAGTAACAGCTCAAGGTACACAGA  
ACAAGAGAGTCTGGTATACAGCAGGTGCTAACAATGCTTGTGGTAGCAAAAGCTATAGCTTTTGGGTGAGAACTCCGA  
35 CCCAAGTCGCGAGTGAAGAGCGAAGGCCCTCTACTCGCCACCGCCCCCGCCCCCAGCTGGGGTCTATAACAGATCACTT  
TCACCTTGGGGAGCCAGAGAGCCCTGGCATCCTAGGTAGCCCCCCCCCGCCCCCCCCCGCAAGCAGCCAGCCCTGCC

[illegible]





[illegible]

[illegible]

[illegible]

CTGTAGCAACCCCTCCGCTGAGGGGCTCCAGGTGGGGGGGUAAGSTGCTSCASTGGGASCTACATGAGAGSTGATGTCTTG  
 GAGTCACCTCGGGTACCATTGTTTAGGGGAGTGGGGATTGTGTGTGGAGACAGGGAGGCTCAAGGATGCTTTTCAACA  
 ATGGTTGATGAGTTGGAACATAACAGGGGGCCATCACACTGGCTCCCATAGCTCTGGGCTTGGCAGCTTCCACATCTGCC  
 CCCCACCCCTGTCTGGCACCAGCTCAAGCTCTGTGATTCTACACATCCCAAAGAGGAAGTAGCCTACTGGGCATGCC  
 5 ACCTCTTCTGGACCATCAGGTGAGAGTGTGGCAAGCCCTAGGCTCCTGTCCAGGATGCAAGGGCTGCCAGATAGGATGCTC  
 AGCTATCTCTGAGCTGGAACTATTTTAGGAATAAGGATTATGCCCGCCCGGGGTTGGCCAGCACCACAGCAGCTGTGC  
 TTGCTAAAGCAAGTGTGTGATTATCTAAAGACAGAGCCCTGGACCCACCCACAGGACAAGTATSTATGCATCTGT  
 TTCATSTATCTGAAAGCGACACAACCATTTTTCACATCATGGCATCTTCTAAGCCCATTTCTTTTTTGTGTGTGT  
 TTGAGACAGGTTTCTCTGTGTAGTCTGGGTGTCTGGAACTCACTTTGTAGACCAGGCTGGGCTCGAAGCTCAGAAATC  
 10 CTGGGATTAAAGGTGTGTGCCACCAGCCCGCCCTAACCCCATTTCTTAATGGTGATCCAGTGGTTGAAATTTGGGGCC  
 ACACACATGTCCATTAGGGATTAGCTGCTGTCTTCTGAGCTACCTGGTACAATCTTTATCCCTGGGGCTGGGCTCCTG  
 ATCCCTGACTCGGGCCCGATCAAGTCCAGTTCTGGGGCCGATCAAGTCCAGTTCTGGGGCCGAACAAGTCCAGTCCCT  
 AGCTCGATTAGCTCATCTGGCTCCCTGGCTGTTCTTACTTACACTCTTCCCTTGTCTGTGGACTTGTGCTTTCTTTA  
 CTCAGTGTGTGCCACAGTCCCTAAGCCACCTCTGTAAGACAACCTAAGATAATACTTCCCTCAAGCACGGAAAGTCTGTG  
 15 AGTCACCACACCCCTCTGGAGGTGTGTGGACACATGTCATGCGTGTGGTTGCGCTTACGTACGTGTGC

Sequence ID No. 18: Human Beer Genomic Sequence (This gene has two exons, at positions 161-427 abd 3186-5219).

20 tagaggagaa gtctttgggg agggtttgct ctgagcacac ccttttcctt ccttcgggg 60  
 ctgagggaaa catgggacca gccctgcccc agcctgtcct cattggctgg catgaagcag 120  
 25 agaggggctt taaaaaggcg accgtgtctc ggctggagac cagagcctgt gctactggaa 180  
 ggtggcgctgc cctcctctgg ctggtaccat gcagctccca ctggccctgt gtctcgtctg 240  
 cctgctggta cacacagcct tccgtgtagt ggagggccag ggggtggcagg cgttcaagaa 300  
 30 tgatgccacg gaaatcatcc ccgagctcgg agagtacccc gagcctccac cggagctgga 360  
 gaacaacaag accatgaacc gggcggagaa cggagggcgg cctccccacc acccctttga 420  
 35 gaccaaaggt atgggggtgga ggagagaatt cttagtaaaa gatcctgggg aggtttttaga 480

aacttctctt tgggaggctt ggaagactgg ggtagacca gtgaagattg ctggcctctg 540  
ccagcactgg tcgaggaaca gtcttgccctg gaggtggggg aagaatggct cgctgggtgca 600  
5 gccttcaaatt tcaggtgcag aggcattgagg caacagacgc tggtagagagc ccagggcagg 660  
gaggacgctg ggggtgggtgag ggtatggcat cagggcatca gaacaggctc aggggctcag 720  
aaaagaaaag gtttcaaaga atctcctcct gggaatatag gagccacgtc cagctgctgg 780  
10 taccactggg aagggaaaca ggtaagggag cctcccatcc acagaacagc acctgtgggg 840  
caccggacac tctatgctgg tgggtggctgt cccaccaca cagaccaca tcatggaatc 900  
15 cccaggaggt gaacccccag ctgaagggg aagaaacagg ttccaggcac tcagtaactt 960  
ggtagtgaga agagctgagg tgtgaacctg gtttgatcca actgcaagat agccctgggtg 1020  
tgtggggggg tgtgggggac agatctccac aaagcagtgg ggaggaaggc cagagaggca 1080  
20 cccctgcagt gtgcattgcc catggcctgc ccaggagct ggcacttgaa ggaatgggag 1140  
ttttcggcac agttttagcc cctgacatgg gtgcagctga gtccaggccc tggaggggag 1200  
25 agcagcatcc tctgtgcagg agtagggaca tctgtcctca gcagccacc cagtcccaac 1260  
cttgccctcat tccaggggag ggagaaggaa gaggaaccct gggttcctgg tcaggcctgc 1320  
acagagaagc ccaggtgaca gtgtgcatct ggctctataa ttggcaggaa tcctgaggcc 1380  
30 atgggggcgt ctgaaatgac acttcagact aagagcttcc ctgtcctctg gccattatcc 1440  
aggtggcaga gaagtccact gccaggtc ctggaccca gccctccccg cctcacaacc 1500  
35 tgttgggact atgggggtgct aaaaagggca actgcatggg aggccagcca ggaccctccg 1560

tcttcaaaat ggaggacaag ggcgcctccc cccacagctc cccttctagg caaggtcagc 1620  
tgggctccag cgactgcctg aagggctgta aggaacccaa acacaaaatg tccaccttgc 1680  
5 tggactccca cgagaggcca cagcccctga ggaagccaca tgctcaaaac aaagtcatga 1740  
tctgcagagg aagtgcctgg cctaggggcg ctattctcga aaagccgcaa aatgccccct 1800  
tccctgggca aatgcccccc tgaccacaca cacattccag ccctgcagag gtgaggatgc 1860  
10 aaaccagccc acagaccaga aagcagcccc agacgatggc agtggccaca tctccccctgc 1920  
tgtgcttgct cttcagagtg ggggtggggg gtggccttct ctgtccctc tctggtttg 1980  
15 tcttaagact atttttcatt ctttcttgct acattggaac tatcccatg aaacctttg 2040  
gggtggactg gtactcacac gacgaccagc tatttaaaaa gctcccaccc atctaagtcc 2100  
accataggag acatggtcaa ggtgtgtgca ggggatcagg ccaggcctcg gagcccaatc 2160  
20 tctgcctgcc cagggagtat caccatgagg cgccattca gataacacag aacaagaaat 2220  
gtgcccagca gagagccagg tcaatgtttg tggcagctga acctgtaggt tttgggtcag 2280  
25 agctcagggc ccctatggta ggaaagtaac gacagtaaaa agcagccctc agtccatcc 2340  
cccagcccag cctcccatgg atgctcgaac gcagagcctc cactcttgcc ggagccaaaa 2400  
ggtgctggga ccccaggga gtggagtccg gagatgcagc ccagcctttt gggcaagttc 2460  
30 ttttctctgg ctgggcctca gtattctcat tgataatgag ggggttgga acactgcctt 2520  
tgattccttt caagtcta atgaattcctgt cctgatcacc tccccttcag tccctgcct 2580  
35 ccacagcagc tgcctgatt tattaccttc aattaacctc tactcctttc tccatccct 2640

gtccaccctt cccaagtggc tggaaaagga atttgggaga agccagagcc aggcagaagg 2700  
tgtgctgagt acttaccctg ccagggccag ggaccctgcg gcacaagtgt ggcttaaatac 2760  
5 ataagaagac cccagaagag aatgataat aataatacat aacagccgac gctttcagct 2820  
atatgtgcca aatggatattt tctgcattgc gtgtgtaatg gattaactcg caatgcttgg 2880  
ggcggcccat tttgcagaca ggaagaagag agaggttaag gaacttgccc aagatgacac 2940  
10 ctgcagtgag cgatggagcc ctgggtgtttg aaccccagca gtcatttggc tccgagggga 3000  
caggggtgcg aggagagctt tccaccagct ctagagcatc tgggaccttc ctgcaataga 3060  
15 tgttcagggg caaaagcctc tggagacagg cttggcaaaa gcagggctgg ggtggagaga 3120  
gacgggcccgg tccagggcag ggggtggccag gcgggcccgc accctcacgc gcgcctctct 3180  
ccacagacgt gtccgagtac agctgccgcg agctgcactt caccgctac gtgaccgatg 3240  
20 ggccgtgccg cagcgccaag ccggtcaccg agctgggtgtg ctccggccag tgcggccccg 3300  
cgcgcttgcg gcccaacgcc atcgccgcg gcaagtgggtg gcgacctagt gggccccgact 3360  
25 tccgctgcat ccccgaccgc taccgcgcgc agcgctgca gctgctgtgt cccgggtggtg 3420  
aggcgccgcg cgcgcgcaag gtgcgcctgg tggcctcgtg caagtgcaag cgctcaccc 3480  
gcttccacaa ccagtcggag ctcaaggact tcgggaccga ggccgctcgg ccgcagaagg 3540  
30 gccggaagcc gcggccccgc gcccgagcg ccaaagccaa ccaggccgag ctggagaacg 3600  
cctactagag cccgcccgcg cccctcccca ccggcgggcg ccccgccct gaaccgcgc 3660  
35 cccacatttc tgctctctgc gcgtgggttg attgtttata tttcattgta aatgcctgca 3720



accagggca gggggctgag accttccagg ccctgaggaa tcccgggagc cggcaaggcc 3780  
ccctcagcc cgccagctga ggggtccac ggggcagggg aggggaattga gagtacaga 3840  
5 cactgagcca cgcagccccg cctctggggc cgcctacctt tgctgggtccc acttcagagg 3900  
aggcagaaat ggaagcattt tcaccgcctt ggggttttaa gggagcgggtg tgggagtggg 3960  
aaagtccagg gactgggttaa gaaagttgga taagattccc cttgcacct cgctgccccat 4020  
10 cagaaagcct gaggcgtgcc cagagcacia gactgggggc aactgtagat gtgggtttcta 4080  
gtcctgggtc tgccactaac ttgctgtgta accttgaact acacaattct ccttcgggac 4140  
15 ctcaatttcc actttgtaaa atgaggggtg aggtgggaat aggatctcga ggagactatt 4200  
ggcatatgat tccaaggact ccagtgcctt ttgaatgggc agaggtgaga gagagagaga 4260  
gaaagagaga gaatgaatgc agttgcattg attcagtgcc aaggtcactt ccagaattca 4320  
20 gagttgtgat gctctcttct gacagccaaa gatgaaaaac aaacagaaaa aaaaaagtaa 4380  
agagtctatt tatggctgac atatttacgg ctgacaaact cctggaagaa gctatgctgc 4440  
25 ttcccagcct ggcttccccg gatgtttggc tacctccacc cctccatctc aaagaaataa 4500  
catcatccat tggggtagaa aaggagaggg tccgaggggtg gtgggagggg tagaaatcac 4560  
atccgccccca acttcccaaa gagcagcatc cctcccccgga cccatagcca tgttttaaag 4620  
30 tcaccttccg aagagaagtg aaaggttcaa ggacactggc cttgcaggcc cgagggagca 4680  
gccatcacia actcacagac cagcacatcc cttttgagac accgccttct gccaccact 4740  
35 cacggacaca tttctgccta gaaaacagct tcttactgct cttacatgtg atggcatatc 4800

ttacactaaa agaattattat tgggggaaaa actacaagtg ctgtacatat gctgagaaac 4860  
tgcagagcat aatagctgcc acccaaaaat ctttttgaaa atcatttcca gacaacctct 4920  
5 tactttctgt gtagttttta attgttaaaa aaaaaaagtt ttaaacagaa gcacatgaca 4980  
tatgaaagcc tgcaggactg gtcgtttttt tggcaattct tccacgtggg acttgtccac 5040  
aagaatgaaa gtagtggttt ttaaagagtt aagttacata tttattttct cacttaagtt 5100  
10 atttatgcaa aagtttttct tgtagagaat gacaatgtta atattgcttt atgaattaac 5160  
agtctgttct tccagagtcc agagacattg ttaataaaga caatgaatca tgaccgaaag 5220  
15 gatgtggtct cattttgtca accacacatg acgtcatttc tgtcaaagtt gacacccttc 5280  
tcttggtcac tagagctcca accttgga cacttttgac tgctctctgg tggcccttgt 5340  
ggcaattatg tcttcctttg aaaagtcag tttatccctt cctttccaaa ccagaccgc 5400  
20 atttcttcac ccagggcatg gtaataacct cagccttgta tccttttagc agcctccct 5460  
ccatgctggc ttccaaaatg ctgttctcat tgtatcactc cctgctcaa aagccttcca 5520  
25 tagctcccc ttgccagga tcaagtgcag tttccctatc tgacatggga ggccttctct 5580  
gcttgactcc cacctcccac tccaccaagc ttcctactga ctccaaatgg tcatgcagat 5640  
ccctgcttcc ttagtttgcc atccacactt agcaccceca ataactaatc ctctttcttt 5700  
30 aggattcaca ttacttgta tctcttcccc taaccttcca gagatgttcc aatctcccat 5760  
gatccctctc tcctctgagg ttccagcccc ttttgtctac accactactt tggttcctaa 5820  
35 ttctgttttc catttgacag tcattcatgg aggaccagcc tggccaagtc ctgcttagta 5880

ctggcataga caacacaaag ccaagtacaa ttcaggacca gctcacagga aacttcatct 5940  
tcttcgaagt gtggatttga tgcctcctgg gtagaaatgt aggatcttca aaagtgggcc 6000  
5 agcctcctgc acttctctca aagtctcgcc tccccagggt gtcttaatag tgctggatgc 6060  
tagctgagtt agcatcttca gatgaagagt aaccctaaag ttactcttca gttgccctaa 6120  
gggtgggatgg tcaactggaa agctttaaat taagtccagc ctaccttggg ggaacccacc 6180  
10 cccacaaaga aagctgaggt cctcctgat gacttgtcag ttaactacc aataaccac 6240  
ttgaattaat catcatcatc aagtctttga taggtgtgag tgggtatcag tggccggtcc 6300  
15 cttcctgggg ctccagcccc cgaggaggcc tcagtgagcc cctgcagaaa atccatgcat 6360  
catgagtgtc tcagggccca gaatatgaga gcaggtagga aacagagaca tcttccatcc 6420  
ctgagaggca gtgcggtcca gtgggtgggg acacgggctc tgggtcaggt ttgtgttgtt 6480  
20 tgtttgtttg ttttgagaca gagtctcgct ctattgccca ggctggagtg cagtgtcaca 6540  
atctcggtt actgcaactt ctgccttccc ggattcaagt gattctcctg cctcagcctc 6600  
25 cagagtagct gggattacag gtgcgtgccca ccacgcctgg ctaatttttg tatttttgat 6660  
agagacgggg tttcaccatg ttggccaggc tagtctcgaa ctcttgacct caagtgatct 6720  
gcctgcctcg gcctcccaaa gtgctgggat tacaggcgtg agccaccaca ccagcccca 6780  
30 ggttggtgtt tgaatctgag gagactgaag caccaagggg taaatgttt tgcccacagc 6840  
catacttggg ctcagttcct tgcctaccc ctcaactgag ctgcttagaa cctggtgggc 6900  
35 acatgggcaa taaccaggct acactgtttt gtaccaagtg ttatgggaat ccaagatagg 6960

agtaatttgc tctgtggagg ggatgagggg tagtggttag ggaaagcttc acaaagtggg 7020  
tggtgcttag agattttcca ggtggagaag ggggcttcta ggcagaaggc atagcccaag 7080  
5 caaagactgc aagtgcattg ctgctcatgg gtagaagaga atccaccatt cctcaacatg 7140  
taccgagtcc ttgccatgtg caaggcaaca tgggggtacc aggaattcca agcaatgtcc 7200  
aaacctaggg tctgctttct gggacctgaa gatacaggat ggatcagccc aggctgcaat 7260  
10 cccattacca cgagggggaa aaaaacctga aggctaaatt gtaggtcggg ttagaggtta 7320  
tttatggaaa gttatattct acctacatgg ggtctataag cctggcgcca atcagaaaag 7380  
15 gaacaaacaa cagacctagc tgggaggggc agcattttgt ttaggggggc ggggcacatg 7440  
ttctgggggt acagccagac tcagggcttg tattaatagt ctgagagtaa gacagacaga 7500  
gggatagaag gaaataggtc cctttctctc tctctctctc tctctctctc actctctctc 7560  
20 tctctcacac acacacacag acacacacac acgctctgta ggggtctact tatgctccaa 7620  
gtacaaatca ggccacattt acacaaggag gtaaaggaaa agaacgttgaggaggccaca 7680  
25 ggaccccaaa attcctgtt ttccttgaat caggcaggac ttacgcagct gggaggggtg 7740  
agagcctgca gaagccacct gcgagtaagc caagttcaga gtcacagaca ccaaagctg 7800  
gtgccatgtc ccacaccgc ccacctcca cctgctcctt gacacagccc tgtgctccac 7860  
30 aaccgggtc ccagatcatt gattatagct ctggggcctg caccgtcctt cctgccacat 7920  
ccccaccca ttcttggaaac ctgccctctg tcttctccct tgtccaaggg caggcaaggg 7980  
35 ctcagctatt gggcagcttt gaccaacagc tgaggctcct tttgtggctg gagatgcagg 8040

aggcagggga atattcctct tagtcaatgc gaccatgtgc ctggtttgcc cagggtggtc 8100  
tcgtttacac ctgtaggcca agcgtaatta ttaacagctc ccacttctac tctaaaaaat 8160  
5 gacccaatct gggcagtaaa ttatatggtg cccatgctat taagagctgc aacttgctgg 8220  
gcgtggtggc tcacacctgt aatcccagta ctttgggacg tcaaggcggg tggatcacct 8280  
gaggtcacga gttagagact ggcctggcca gcatggcaaa accccatctt tactaaaaat 8340  
10 acaaaaatta gcaaggcatg gtggcatgca cctgtaatcc caggtactcg ggaggctgag 8400  
acaggagaat ggcttgaacc caggaggcag aggttgcagt gagccaagat tgtgccactg 8460  
15 ccctccagcc ctggcaacag agcaagactt catctcaaaa gaaaaaggat actgtcaatc 8520  
actgcaggaa gaaccaggt aatgaatgag gagaagagag gggctgagtc accatagtgg 8580  
cagcaccgac tcctgcagga aaggcgagac actgggtcat ggggtactgaa ggggtgccctg 8640  
20 aatgacgttc tgctttagag accgaacctg agccctgaaa gtgcatgcct gttcatgggt 8700  
gagagactaa attcatcatt ccttggcagg tactgaatcc tttcttacgg ctgccctcca 8760  
25 atgccaatt tccctacaat tgtctggggt gcctaagctt ctgcccacca agagggccag 8820  
agctggcagc gagcagctgc aggtaggaga gataggtacc cataagggag gtgggaaaga 8880  
gagatggaag gagaggggtg cagagcacac acctcccctg cctgacaact tcctgagggc 8940  
30 tggatcatgcc agcagattta aggcggaggc aggggagatg gggcgggaga ggaagtgaaa 9000  
aaggagaggg tggggatgga gaggaagaga gggatgatcat tcattcattc cattgctact 9060  
35 gactggatgc cagctgtgag ccaggcacca ccctagctct gggcatgtgg ttgtaatctt 9120

ggagcctcat ggagctcaca gggagtgcctg gcaaggagat ggataatgga cggataacaa 9180  
ataaacattt agtacaatgt ccgggaatgg aaagttctcg aaagaaaaat aaagctggtg 9240  
5 agcatataga cagccctgaa ggcggccagg ccaggcattt ctgaggaggt ggcatttgag 9300  
c 9301

- 10 From the foregoing, it will be appreciated that, although specific embodiments of the invention have been described herein for purposes of illustration, various modifications may be made without deviating from the spirit and scope of the invention. Accordingly, the invention is not limited except as by the appended claims.

CLAIMS

We claim:

1. An isolated nucleic acid molecule selected from the group consisting of:
  - 5 (a) an isolated nucleic acid molecule comprising sequence ID Nos., 1, 5, 9, 11, 13, or, 15, or complementary sequence thereof;
  - (b) an isolated nucleic acid molecule that specifically hybridizes to the nucleic acid molecule of (a) under conditions of high stringency; and
  - (c) an isolated nucleic acid that encodes a TGF-beta binding-protein  
10 according to (a) or (b).
2. The isolated nucleic acid molecule according to claim 1 wherein said nucleic acid molecule encodes a protein comprising the protein of Sequence ID NO. 2.
3. The isolated nucleic acid molecule according to claim 1 wherein  
15 said nucleic acid molecule encodes a protein comprising the protein of Sequence ID NO. 6.
4. The isolated nucleic acid molecule according to claim 1 wherein said nucleic acid molecule encodes a protein comprising the protein of Sequence ID NO. 10.
- 20 5. The isolated nucleic acid molecule according to claim 1 wherein said nucleic acid molecule encodes a protein comprising the protein of Sequence ID NO. 12.
6. The isolated nucleic acid molecule according to claim 1 wherein said nucleic acid molecule encodes a protein comprising the protein of Sequence ID  
25 NO. 14.
7. The isolated nucleic acid molecule according to claim 1 wherein said nucleic acid molecule encodes a protein comprising the protein of Sequence ID NO. 16.

8. An expression vector, comprising a promoter operably linked to a nucleic acid molecule according to any one of claims 1 to 7.

9. The expression vector according to claim 8 wherein said promoter is selected from the group consisting of CMV I-E promoter, SV40 early promoter and MuLV LTR.

10. The expression vector according to claim 8 wherein said promoter is a tissue-specific promoter.

11. A method of producing a TGF-beta binding protein, comprising, culturing a cell which contains a vector according to claim 8 under conditions and for a time sufficient to produce said protein.

12. The method according to claim 11, further comprising the step of purifying said protein.

13. A viral vector capable of directing the expression of a nucleic acid molecule according to any one of claims 1 to 7.

14. The viral vector according to claim 13 wherein said vector is selected from the group consisting of herpes simplex viral vectors, adenoviral vectors, adenovirus-associated viral vectors and retroviral vectors.

15. A host cell carrying a vector according to any one of claims 8 to 14.

16. The host cell according to claim 15 wherein said cell is selected from the group consisting of a human cell, dog cell, monkey cell, rat cell and mouse cell.

17. An isolated protein, comprising a TGF-beta binding-protein encoded by the nucleic acid molecule according to any one of claims 1 to 7.

18. An antibody which specifically binds to the protein according to claim 17.



19. The antibody according to claim 18 wherein said antibody is a monoclonal antibody.

20. The antibody according to claim 19 wherein said monoclonal antibody is a murine or human antibody.

5 21. The antibody according to claim 18 wherein said antibody is selected from the group consisting of  $F(ab')_2$ ,  $F(ab)_2$ ,  $Fab'$ ,  $Fab$ , and  $Fv$ .

22. A hybridoma which produces an antibody according to claim 19.

10 23. A fusion protein, comprising a first polypeptide segment comprising a TGF-beta binding-protein encoded by the nucleic acid molecule according to any one of claims 1 to 7, or a portion thereof of at least 10 amino acids in length, and a second polypeptide segment comprising a non-TGF-beta binding-protein.

24. The fusion protein according to claim 23 wherein said first polypeptide segment is at least 20 amino acids in length.

15 25. The fusion protein according to claim 23 wherein said first polypeptide segment is at least 50 amino acids in length.

26. The fusion protein according to claim 23 wherein said second polypeptide comprises multiple anionic amino acid residues.

20 27. An isolated oligonucleotide which hybridizes to a nucleic acid molecule according to Sequence ID NOs. 1, 3, 5, 7, 9, 11, 13, or 15, or the complement thereto, under conditions of high stringency.

28. The isolated oligonucleotide according to claim 27 wherein said oligonucleotide is at least 20 nucleotides in length.

29. The isolated oligonucleotide according to claim 27 wherein said oligonucleotide is at least 30 nucleotides in length.

25 30. The isolated oligonucleotide according to claim 27 wherein said

oligonucleotide is at least 50 nucleotides in length.

31. The isolated oligonucleotide according to claim 27 wherein said oligonucleotide is between 50 to 100 nucleotides in length.

32. A pair of primers which specifically amplifies all or a portion of  
5 a nucleic acid molecule according to any one of claims 1 to 7.

33. A ribozyme which cleaves RNA encoding a protein according to claim 17.

34. The ribozyme according to claim 33 wherein said protein comprises the protein of Sequence ID NO. 2.

10 35. The ribozyme according to claim 33 wherein said protein comprises the protein of Sequence ID NO. 6.

36. The ribozyme according to claim 33 wherein said RNA encodes a protein comprising the protein of Sequence ID NO. 10.

15 37. The ribozyme according to claim 33 wherein said RNA encodes a protein comprising the protein of Sequence ID NO. 12.

38. The ribozyme according to claim 33 wherein said RNA encodes a protein comprising the protein of Sequence ID NO. 14.

39. The ribozyme according to claim 33 wherein said RNA encodes a protein comprising the protein of Sequence ID NO. 16.

20 40. The ribozyme according to claim 33 wherein said ribozyme is composed of ribonucleic acids.

41. The ribozyme according to claim 40 wherein one or more of said ribonucleic acids are 2'-O-methyl ribonucleic acids.

42. The ribozyme according to claim 33 wherein said ribozyme is

composed of a mixture of deoxyribonucleic acids and ribonucleic acids.

43. The ribozyme according to claim 33 wherein said ribozyme is composed of nucleic acids having phosphothioate linkages.

44. A nucleic acid molecule comprising a nucleic acid sequence  
5 which encodes a ribozyme according to claim 33.

45. The nucleic acid molecule of claim 44, wherein the nucleic acid is DNA or cDNA.

46. The nucleic acid molecule of claim 44, under the control of a promoter to transcribe the nucleic acid.

10 47. A host cell comprising the ribozyme of claim 33.

48. A vector, comprising the nucleic acid molecule of claim 44.

49. The vector of claim 54, wherein the vector is a plasmid, a virus, retrotransposon or a cosmid.

50. The vector of claim 49 wherein said virus is selected from the  
15 group consisting of retroviruses, adenoviruses, and adeno-associated viruses.

51. A host cell containing the vector according to any one of claims 48 to 50.

52. The host cell according to claim 51 wherein said host cell is stably transformed with said vector.

20 53. The host cell according to claim 51 wherein the host cell is a human cell.

54. A method for producing a ribozyme, comprising providing DNA encoding the ribozyme according to claim 33 under the transcriptional control of a promoter, and transcribing the DNA to produce the ribozyme.

55. The method of claim 54 wherein the ribozyme is produced *in vitro*.
56. The method of claim 54, further comprising purifying the ribozyme.
- 5 57. A method for increasing bone mineralization, comprising introducing into a warm-blooded animal an effective amount of the ribozyme according to any one of claims 33 to 43.
58. A method of increasing bone mineralization, comprising introducing into a patient an effective amount of the nucleic acid molecule of claim 44,  
10 under conditions favoring transcription of the nucleic acid molecule to produce a ribozyme.
59. A pharmaceutical composition, comprising the ribozyme according to any one of claims 33 to 43, and a pharmaceutically acceptable carrier or diluent.
- 15 60. A pair of primers capable of specifically amplifying all or a portion of a nucleic acid molecule according to any one claims 1 to 7.
61. A method for detecting a nucleic acid molecule which encodes a TGF-beta binding protein, comprising incubating an oligonucleotide according to any one of claims 27 to 31 under conditions of high stringency, and detecting hybridization  
20 of said oligonucleotide.
62. The method according to claim 61 wherein said oligonucleotide is labeled.
63. The method according to claim 61 wherein said oligonucleotide is bound to a solid support.
- 25 64. A method for detecting a TGF-beta binding protein, comprising incubating an antibody according to any one of claims 18 to 21 under conditions and for a time sufficient to permit said antibody to bind to a TGF-beta binding protein, and

detecting said binding.

65. The method according to claim 64 wherein said antibody is bound to a solid support.

66. The method according to claim 64 wherein said antibody is  
5 labeled.

67. The method according to claim 66 wherein said antibody is labeled with a marker selected from the group consisting of enzymes, fluorescent proteins, and radioisotopes.

68. A transgenic animal whose germ cells and somatic cells contain a  
10 nucleic acid molecule encoding a TGF-beta binding-protein according to claim 1 which is operably linked to a promoter effective for the expression of said gene, said gene being introduced into said animal, or an ancestor of said animal, at an embryonic stage, with the proviso that said animal is not a human.

69. The transgenic animal according to claim 68 wherein TGF-beta  
15 binding-protein is expressed from a vector according to any one of claims 8 to 10.

70. A transgenic knockout animal, comprising an animal whose germ cells and somatic cells comprise a disruption of at least one allele of an endogenous nucleic acid molecule which hybridizes to the nucleic acid molecule according to claim 1, wherein said disruption prevents transcription of messenger RNA  
20 from said allele as compared to an animal without said disruption, with the proviso that said animal is not a human.

71. The transgenic animal according to claim 70 wherein said disruption is a nucleic acid deletion, substitution, or, insertion.

72. The transgenic animal according to claim 68 or 70 wherein the  
25 animal is selected from the group consisting of a mouse, a rat and a dog.

73. A method for determining whether a candidate molecule is capable of increasing bone mineral content, comprising:

(a) mixing one or more candidate molecules with TGF-beta-binding-protein encoded by the nucleic acid molecule according to any one of claims 1 to 7 and  
5 a selected member of the TGF-beta family of proteins;

(b) determining whether the candidate molecule alters the signaling of the TGF-beta family member, or alters the binding of the TGF-beta binding-protein to the TGF-beta family member.

74. The method according to claim 73 wherein said member of the  
10 TGF-beta family of proteins is BMP6.

75. A method for determining whether a candidate molecule is capable of increasing bone mineral content, comprising: determining whether a candidate molecule inhibits the binding of TGF-beta binding-protein to bone, or an analogue thereof.

15 76. The method according to claim 75 wherein said analogue of bone is hydroxyapatite.

77. A kit for detection of TGF-beta binding-protein gene expression, comprising a container that comprises a nucleic acid molecule, wherein said nucleic acid molecule is selected from the group consisting of (a) a nucleic acid molecule  
20 comprising the nucleotide sequence of SEQ ID NO: 1, 5, 7, 9, 11, 13, or 15; (b) a nucleic acid molecule comprising the complement of the nucleotide sequence of (a); (c) a nucleic acid molecule that is a fragment of (a) or (b) of at least 20 nucleotides in length.

78. A kit for detection of TGF-beta binding-protein,  
25 comprising a container that comprises an antibody according to any one of claims 18 to 21.

79. An antisense oligonucleotide, comprising a nucleic acid molecule which hybridizes to a nucleic acid molecule according to Sequence ID NOs. 1, 3, 5, 7, 9, 11, 13, or 15, or the complement thereto, and wherein  
30 said oligonucleotide inhibits the expression of TGF-beta binding protein according to

claim 17.

80. The oligonucleotide according to claim 79 wherein said oligonucleotide is 15 nucleotides in length.

81. The oligonucleotide according to claim 79  
5 wherein said oligonucleotide is 20 nucleotides in length.

82. The oligonucleotide according to claim 79 wherein said oligonucleotide is 50 nucleotides in length.

83. The oligonucleotide according to claim 79, wherein said oligonucleotide is comprised of one or more nucleic acid analogs.

84. The oligonucleotide according to claim 79,  
10 wherein said oligonucleotide is comprised of one or more ribonucleic acids.

85. The oligonucleotide according to claim 79, wherein said oligonucleotide is comprised of one or more deoxyribonucleic acids.

86. The oligonucleotide according to claim 79  
15 wherein said oligonucleotide sequence comprises one or more modified covalent linkages.

87. The oligonucleotide according to claim 86 wherein said modified covalent linkage is selected from the group consisting of a phosphorothioate linkage, a phosphotriester linkage, a methyl phosphonate linkage, a  
20 methylene(methylimino) linkage, a morpholino linkage, an amide linkage, a polyamide linkage, a short chain alkyl intersugar linkage, a cycloalkyl intersugar linkage, a short chain heteroatomic intersugar linkage and a heterocyclic intersugar linkage.

## Common Cysteine Backbone

1	human_gremlin.pro	-----	-----	-----	50
	human_cerberus.pro	MHLLLFQLLV	LLPLGKTTRH	QDGRQNQSSL	SPVLLPRNQR ELPTGNHEEA
	human_dan.pro	-----	-----	-----	-----
	human_beer.pro	-----	-----	-----	-----
		51			100
	human_gremlin.pro	-----	-----M	SRTAYTVGAL	LLLLGTLLPA AEGKKKGSQG
	human_cerberus.pro	EEKPDLFVAV	PHLVAT.SPA	GEGQRQREKM	LSRFGRFWKK PEREMHPSRD
	human_dan.pro	-----	-----	-----	-----
	human_beer.pro	-----	-----	-----	-----MQLPLA LCLVCLLVHT
		101			150
	human_gremlin.pro	AI.PPPDKAQ	HNDSEQTQSP	QQPGSRNRGR	GQGRGTAMPG EEVLESSQEA
	human_cerberus.pro	SDSEPFPPGT	QSLIQPID.G	MKMEKSPLRE	EAKKFWHHFM FRKTPASQGV
	human_dan.pro	-----	-----	-----	MLRVLVGAVL PAMLLAAPP
	human_beer.pro	AFRVVEGQGW	QAFKNDATETI	IPELGEYPEP	PPELENNKTM NRAENGGRPP
		151	↓	↓	↓ ↓ 200
	human_gremlin.pro	LHVTERKYLK	RDWCKTOPLK	QTIHEEGCNS	RTIINRF.CY GQCNSFYIPR
	human_cerberus.pro	ILPIKSHEVH	WETCRTVPFS	QTITHEGCEK	VVVQNNL.CF GKCGSVHFP.
	human_dan.pro	INKLALFPDK	SAWCEAKNIT	QIVGHSGCEA	KSIQNRA.CL GQCFSYSVPN
	human_beer.pro	HHPFETKDVS	EYSCRELHFT	RYVTDGPCRS	AKPVTTELVC S GQCGPARLLP
		201	↓	↓	250
	human_gremlin.pro	HIRKEEGSFQ	SCSF...CKP	KKFTTMMVTL	NCPELQPPTK K.KRVTRVKQ
	human_cerberus.pro	..GAAQHSHT	SCSH...CLP	AKFTTMHLPL	NCTELSSVIK V...VMLVEE
	human_dan.pro	TFPQSTESLV	HCDS...CMP	AQSMWEIVTL	ECPGHEEVPR VDKLVEKILH
	human_beer.pro	NAIGRGKWWR	PSGPDFRCIP	DRYRAQRVQL	LCPGGEAPRA RKVRLVAS..
		↓ ↓			300
	human_gremlin.pro	CRC.ISIDLD	-----	-----	-----
	human_cerberus.pro	CQCKVKTEHE	DGHILHAGSQ	DSFIPGVSA-	-----
	human_dan.pro	CSCQACGKEP	SHEGLSVYVQ	GEDGPGSQPG	THPHPHPHPH PGGQTPEPED
	human_beer.pro	CKCKRLTRFH	NQSELKDFGT	EAARPQKGRK	PRPRARSAKA NQAELENAY-
		301	314		
	human_gremlin.pro	-----	-----		
	human_cerberus.pro	-----	-----		
	human_dan.pro	PPGAPHTEEE	GAED		
	human_beer.pro	-----	-----		

Figure 1



# Human Beer Gene Expression by RT-PCR

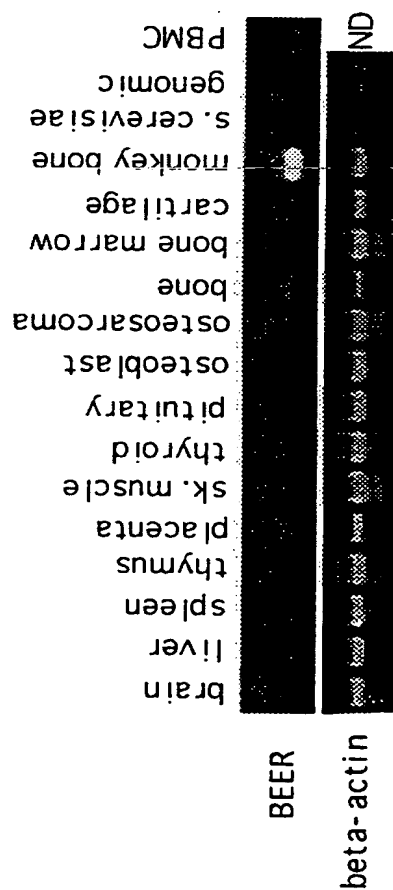


Fig. 2

# RNA In Situ Hybridization of Mouse Embryo Sections

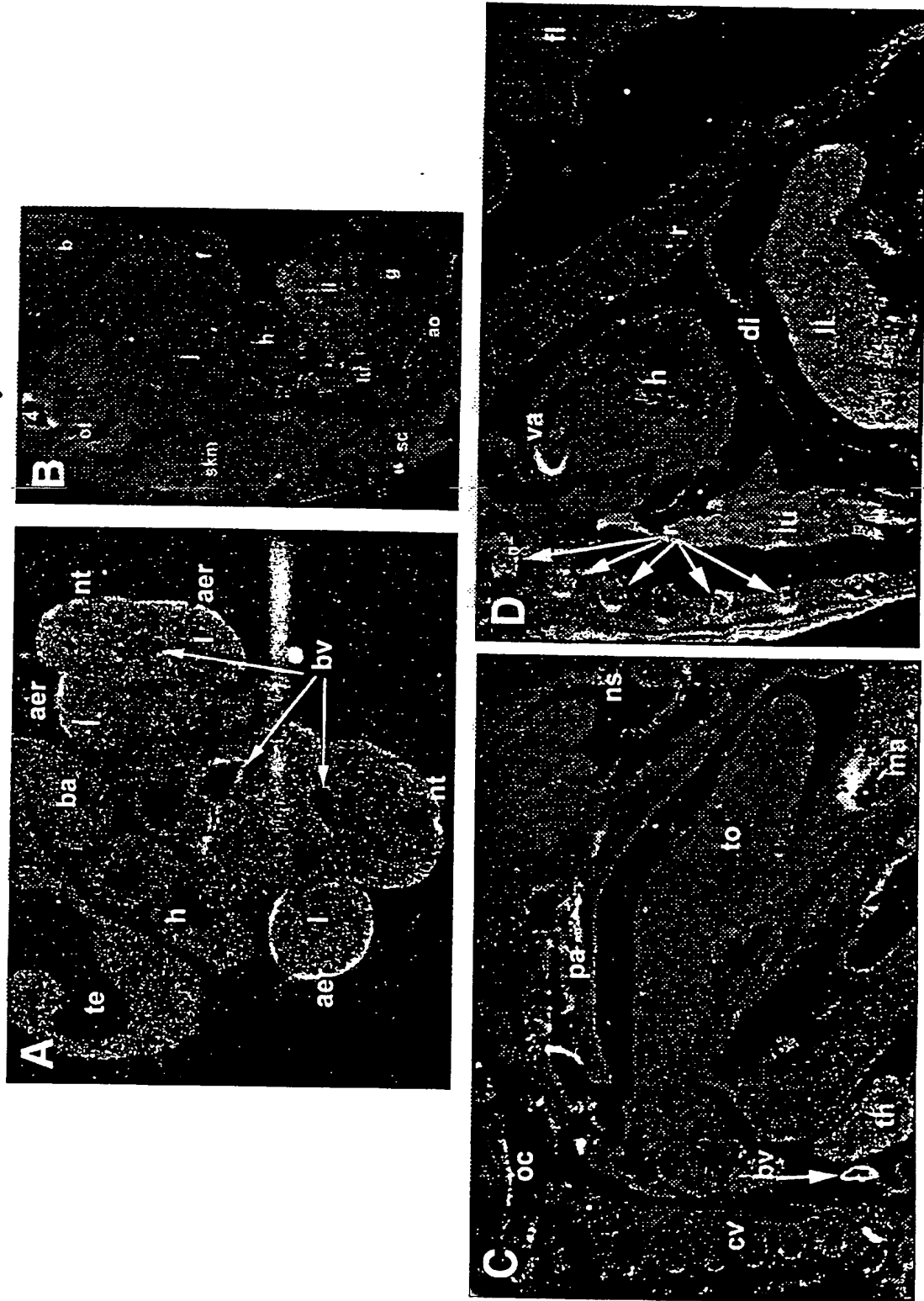


Fig. 3

# Antibody Selectivity

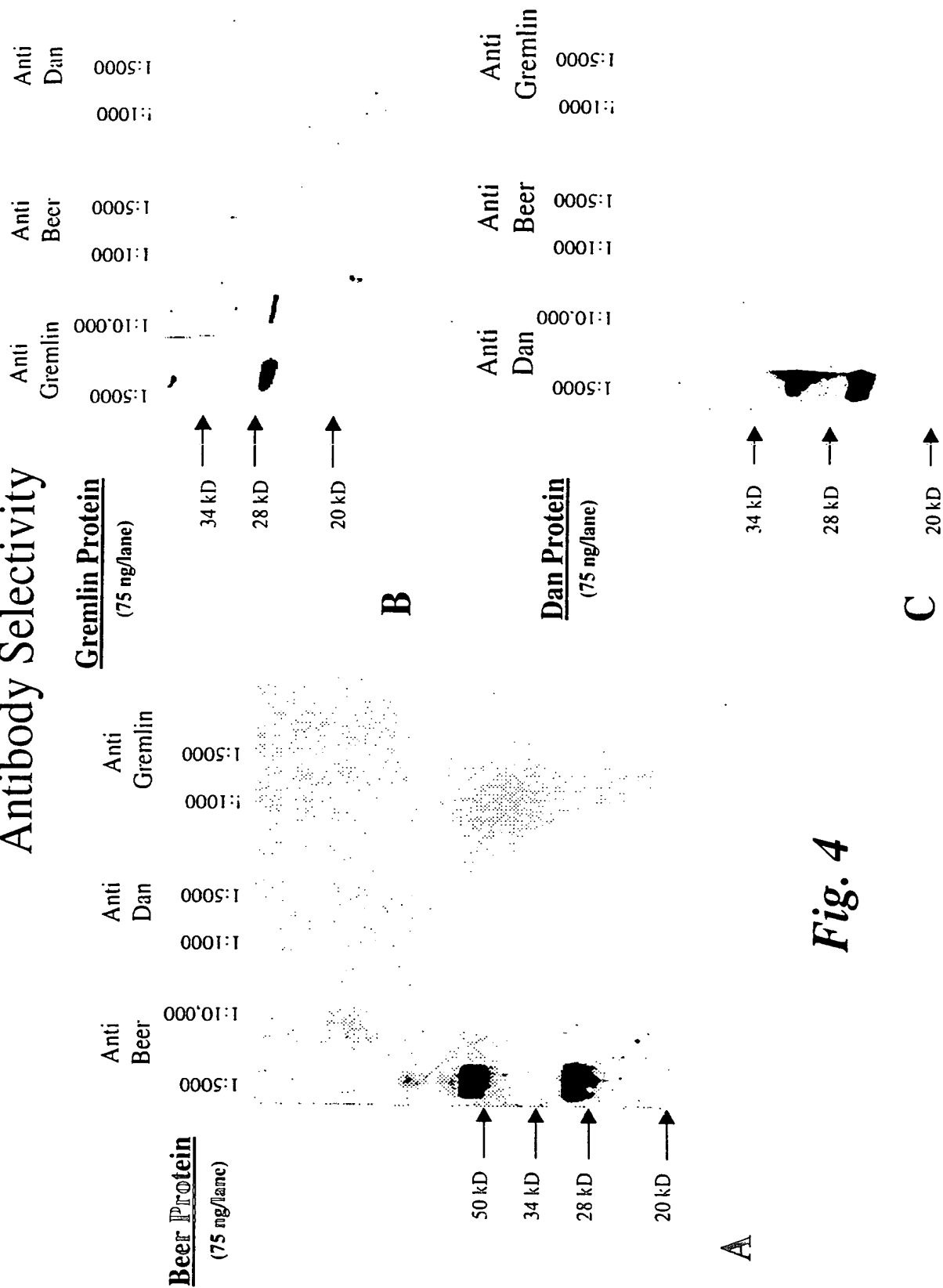
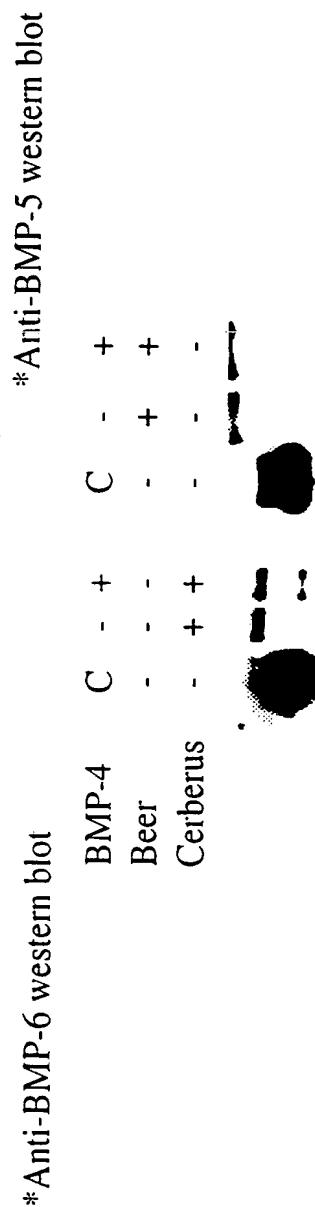
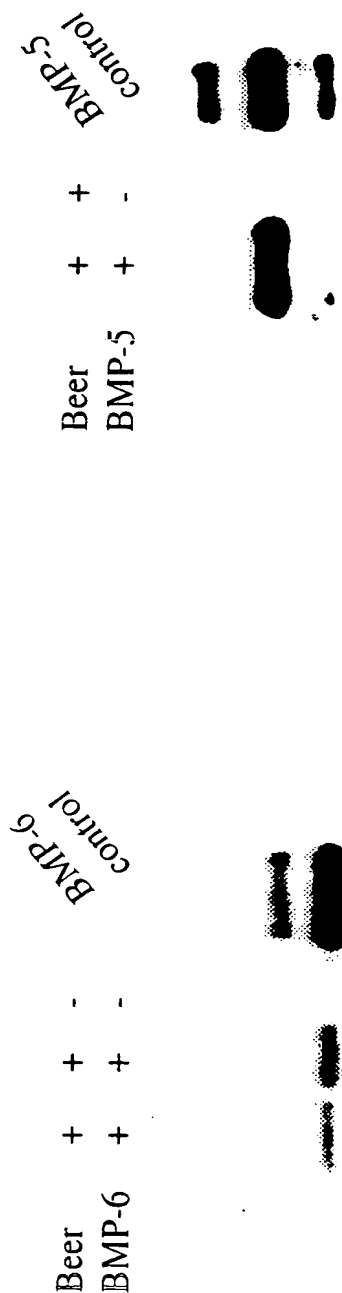


Fig. 4

## Evaluation of Beer binding to BMP family members Anti-FLAG Immunoprecipitation



**\*\*Anti-BMP-4 western blot**

**Fig. 5**

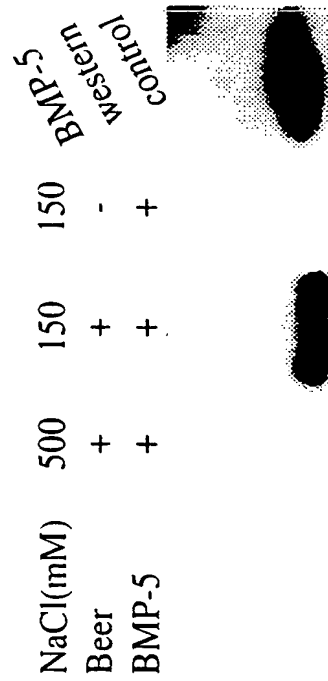
# BMP-5/Beer Dissociation Constant Characterization

.75 1.5 7.5 15 30 60 120 nM BMP-5



\*Anti-FLAG immunoprecipitation \*Anti-BMP-5 western blot

## Ionic Disruption of BMP-5/ Beer Binding



\* Anti FLAG immunoprecipitation  
\* Anti BMP-5 western

Fig. 6

## SEQUENCE LISTING

<110> Brunkow, Mary E.  
5 Galas, David J.  
Kovacevich, Brian  
Mulligan, John T.  
Paeper, Bryan W.  
Van Ness, Jeffrey  
10 Winkler, David G.

<120> COMPOSITIONS AND METHODS FOR INCREASING  
BONE MINERALIZATION  
15

<130> 240083.508

<140> US  
<141> 1999-11-24  
20

<160> 41

<170> FastSEQ for Windows Version 3.0

25 <210> 1  
<211> 2301  
<212> DNA  
<213> Homo sapien

30 <400> 1

agagcctgtg ctactggaag gtggcgtgcc ctccctctggc tggtagcatg cagctccac 60  
tggccctgtg tctcgtctgc ctgctggtac acacagcctt ccgtgtagtg gagggccagg 120  
ggtggcaggc gttcaagaat gatgccacgg aaatcatccc cgagctcggg gagtaccaccg 180  
agcctccacc ggagctggag aacaacaaga ccatgaaccg ggcggagaac ggagggcggc 240  
35 ctccccacca cccctttgag accaaagacg tgtccgagta cagctgccgc gagctgcact 300  
tcaccgcta cgtgaccgat gggccgtgcc gcagcgccaa gccggtcacc gagctggtgt 360

gctccggcca gtgcggcccg gcgcgcctgc tgcccaacgc catcggccgc ggcaagtggg 420  
 ggcgacctag tgggcccgcac ttccgctgca tccccgaccg ctaccgcgcg cagcgcgtgc 480  
 agctgctgtg tcccgggtggg gaggcgccgc gcgcgcgcaa ggtgcgcctg gtggcctcgt 540  
 gcaagtgcaa gcgcctcacc cgcttccaca accagtgcga gctcaaggac ttcgggaccg 600  
 5 agggcgctcg gccgcagaag ggccggaagc cgcggccccc cgcccggagc gccaaagcca 660  
 accaggccga gctggagaac gcctactaga gcccgcgccg gccctcccc accggcgggc 720  
 gccccggccc tgaaccgcgc cccacattt ctgtcctctg cgcgtgggtt gattgtttat 780  
 atttcattgt aaatgcctgc aaccaggggc agggggctga gaccttccag gccctgagga 840  
 atccccggcg ccggcaaggc cccctcagc ccgccagctg aggggtccca cggggcaggg 900  
 10 gagggaattg agagtcacag aactgagcc acgcagcccc gcctctgggg ccgcctacct 960  
 ttgctgggtcc cacttcagag gaggcagaaa tggaaagcatt ttcaccgcc tggggtttta 1020  
 agggagcggg gtgggagtg gaaagtccag ggactgggta agaaagttgg ataagattcc 1080  
 cccttgccacc tcgctgccc tcaaaaagcc tgaggcgtgc ccagagcaca agactggggg 1140  
 caactgtaga tgtgggtttt agtcctgggt ctgccactaa cttgctgtgt aaccttgaac 1200  
 15 tacacaattc tccttcggga cctcaatttc cactttgtaa aatgaggggt gaggtgggaa 1260  
 taggatctcg aggagactat tggcatatga ttccaaggac tccagtgcct tttgaatggg 1320  
 cagaggtgag agagagagag agaaagagag agaataagtg cagttgcatt gattcagtgc 1380  
 caaggtcact tccagaattc agagttgtga tgctctcttc tgacagccaa agatgaaaaa 1440  
 caaacagaaa aaaaaagta aagagtctat ttatggctga catatttacg gctgacaaac 1500  
 20 tcctggaaga agctatgctg cttcccagcc tggcttcccc ggatgtttgg ctacctccac 1560  
 ccctccatct caaagaaata acatcatcca ttggggtaga aaaggagagg gtccgagggg 1620  
 ggtgggaggg atagaaatca catccgcccc aacttcccaa agagcagcat cctcccccg 1680  
 acccatagcc atgttttaaa gtcaccttcc gaagagaagt gaaaggttca aggacactgg 1740  
 ccttgaggc ccgagggagc agccatcaca aactcacaga ccagcacatc ccttttgaga 1800  
 25 caccgccttc tgcccaccac tcacggacac atttctgcct agaaaacagc ttcttactgc 1860  
 tcttacatgt gatggcatat cttacactaa aagaatatta ttgggggaaa aactacaagt 1920  
 gctgtacata tgctgagaaa ctgcagagca taatagctgc caccacaaaa tctttttgaa 1980  
 aatcatttcc agacaacctc ttactttctg ttagtttttt aattgttaaa aaaaaaagt 2040  
 tttaaacaga agcacatgac atatgaaagc ctgcaggact ggtcgttttt ttggcaattc 2100  
 30 ttccacgtgg gacttgcca caagaatgaa agtagtgggt tttaaagagt taagttacat 2160  
 atttatttcc tcacttaagt tatttatgca aaagttttcc ttgtagagaa tgacaatggt 2220  
 aatattgctt tatgaattaa cagtctgttc ttccagagtc cagagacatt gttaataaag 2280  
 acaatgaatc atgaccgaaa g 2301

35 <210> 2

<211> 213

&lt;212&gt; PRT

&lt;213&gt; Homo sapien

&lt;400&gt; 2

5 Met Gln Leu Pro Leu Ala Leu Cys Leu Val Cys Leu Leu Val His Thr  
 1 5 10 15  
 Ala Phe Arg Val Val Glu Gly Gln Gly Trp Gln Ala Phe Lys Asn Asp  
 20 25 30  
 Ala Thr Glu Ile Ile Pro Glu Leu Gly Glu Tyr Pro Glu Pro Pro Pro  
 10 35 40 45  
 Glu Leu Glu Asn Asn Lys Thr Met Asn Arg Ala Glu Asn Gly Gly Arg  
 50 55 60  
 Pro Pro His His Pro Phe Glu Thr Lys Asp Val Ser Glu Tyr Ser Cys  
 65 70 75 80  
 15 Arg Glu Leu His Phe Thr Arg Tyr Val Thr Asp Gly Pro Cys Arg Ser  
 85 90 95  
 Ala Lys Pro Val Thr Glu Leu Val Cys Ser Gly Gln Cys Gly Pro Ala  
 100 105 110  
 Arg Leu Leu Pro Asn Ala Ile Gly Arg Gly Lys Trp Trp Arg Pro Ser  
 20 115 120 125  
 Gly Pro Asp Phe Arg Cys Ile Pro Asp Arg Tyr Arg Ala Gln Arg Val  
 130 135 140  
 Gln Leu Leu Cys Pro Gly Gly Glu Ala Pro Arg Ala Arg Lys Val Arg  
 145 150 155 160  
 25 Leu Val Ala Ser Cys Lys Cys Lys Arg Leu Thr Arg Phe His Asn Gln  
 165 170 175  
 Ser Glu Leu Lys Asp Phe Gly Thr Glu Ala Ala Arg Pro Gln Lys Gly  
 180 185 190  
 Arg Lys Pro Arg Pro Arg Ala Arg Ser Ala Lys Ala Asn Gln Ala Glu  
 30 195 200 205  
 Leu Glu Asn Ala Tyr  
 210

&lt;210&gt; 3

35 &lt;211&gt; 2301

&lt;212&gt; DNA



&lt;213&gt; Homo sapien

&lt;400&gt; 3

	agagcctgtg	ctactggaag	gtggcgtgcc	ctcctctggc	tggtaccatg	cagctcccac	60
5	tgggcctgtg	tctcgtctgc	ctgctggtac	acacagcctt	ccgtgtagtg	gagggctagg	120
	ggtggcaggc	gttcaagaat	gatgccacgg	aaatcatccc	cgagctcgga	gagtaccccg	180
	agcctccacc	ggagctggag	aacaacaaga	ccatgaaccg	ggcggagaac	ggagggcggc	240
	ctccccacca	cccctttgag	accaaagacg	tgtccgagta	cagctgccgc	gagctgcact	300
	tcacccgcta	cgtgaccgat	gggccgtgcc	gcagcgccaa	gccggtcacc	gagctggtgt	360
10	gctccggcca	gtgcggcccc	gcgcgcctgc	tgcccaacgc	catcggccgc	ggcaagtggg	420
	ggcgacctag	tgggccccgac	ttccgctgca	tccccgaccg	ctaccgcgcg	cagcgcgtgc	480
	agctgctgtg	tcccgggtgt	gagggcggcg	gcgcgcgcaa	gggtgcgcctg	gtggcctcgt	540
	gcaagtgcaa	gcgcctcacc	cgcttcacac	accagtcgga	gctcaaggac	ttcgggaccg	600
	agggcgctcg	gccgcagaag	ggccggaagc	cgcgggccccg	cgccccggagc	gccaaagcca	660
15	accaggccga	gctgggagaac	gcctactaga	gccccgccgc	gccccctccc	accggcgggc	720
	gccccggccc	tgaacccgcg	ccccacattt	ctgtcctctg	cgcgtggttt	gattgtttat	780
	atctcattgt	aaatgcctgc	aaccagggc	agggggctga	gaccttccag	gccctgagga	840
	atccccggcg	ccggcaaggc	ccccctcagc	ccgccagctg	aggggtccca	cggggcaggg	900
	gaggggaattg	agagtcacag	acactgagcc	acgcagcccc	gcctctgggg	ccgcctacct	960
20	ttgctggtcc	cacttcagag	gaggcagaaa	tggaagcatt	ttcaccgccc	tggggtttta	1020
	agggagcggg	gtgggagtgg	gaaagtccag	ggactgggta	agaaagttag	ataagattcc	1080
	cccttgccacc	tcgctgcccc	tcagaaaagc	tgaggcgtgc	ccagagcaca	agactggggg	1140
	caactgtaga	tgtgggtttct	agtcctggct	ctgccactaa	cttgctgtgt	aaccttgaac	1200
	tacacaattc	tccttcggga	cctcaatttc	cactttgtaa	aatgaggggtg	gaggtgggaa	1260
25	taggatctcg	aggagactat	tggcatatga	ttccaaggac	tccagtgcct	tttgaatggg	1320
	cagagggtgag	agagagagag	agaaagagag	agaatgaatg	cagttgcatt	gattcagtgc	1380
	caaggctcact	tccagaattc	agagttgtga	tgtctctctc	tgacagccaa	agatgaaaaa	1440
	caaacagaaa	aaaaaaaagta	aagagtctat	ttatggctga	catatttacg	gctgacaaac	1500
	tcctggaaga	agctatgctg	cttcccagcc	tggcttcccc	ggatgtttgg	ctacctccac	1560
30	ccctccatct	caaagaaata	acatcatcca	ttggggtaga	aaaggagagg	gtccgagggt	1620
	gggtgggaggg	atagaaatca	catccgcccc	aacttcccaa	agagcagcat	ccctcccccg	1680
	acccatagcc	atgtttttaa	gtcaccttcc	gaagagaagt	gaaagggttca	aggacactgg	1740
	ccttgccaggc	ccgagggagc	agccatcaca	aactcacaga	ccagcacatc	ccttttgaga	1800
	caccgccttc	tgccaccac	tcacggacac	atctctgcct	agaaaacagc	ttcttactgc	1860
35	tcttacatgt	gatggcatat	cttacactaa	aagaatatta	ttgggggaaa	aactacaagt	1920
	gctgtacata	tgctgagaaa	ctgcagagca	taatagctgc	cacccaaaaa	tctttttgaa	1980

aatcatttcc agacaacctc ttactttctg ttagtatttt aattgttaaa aaaaaaaagt 2040  
 tttaaacaga agcacatgac atatgaaagc ctgcaggact ggctcgtttt ttggcaattc 2100  
 ttccacgtgg gacttgcca caagaatgaa agtagtggtt tttaaagagt taagttacat 2160  
 atttattttc tactttaagt tatttatgca aaagtttttc ttgtagagaa tgacaatgtt 2220  
 5 aatattgctt tatgaattaa cagtctgttc ttccagagtc cagagacatt gttaataaag 2280  
 acaatgaatc atgaccgaaa g 2301

<210> 4

<211> 23

10 <212> PRT

<213> Homo sapien

<400> 4

Met Gln Leu Pro Leu Ala Leu Cys Leu Val Cys Leu Leu Val His Thr  
 15 1 5 10 15  
 Ala Phe Arg Val Val Glu Gly  
 20

<210> 5

20 <211> 2301

<212> DNA

<213> Homo sapien

<400> 5

25 agagcctgtg ctactggaag gtggcgtgcc ctccctctggc tggtagcatg cagctccac 60  
 tggccctgtg tctcatctgc ctgctggtac acacagcctt ccgtgtagtg gagggccagg 120  
 ggtggcaggc gttcaagaat gatgccacgg aaatcatccg cgagctcggg gagtaccctc 180  
 agcctccacc ggagctggag aacaacaaga ccatgaaccg ggcggagAAC ggagggcggc 240  
 ctccccacca cccctttgag accaaagacg tgtccgagta cagctgccgc gagctgcact 300  
 30 tccccgcta cgtgaccgat gggcgtgcc gcagcgccaa gccggtcacc gagctggtgt 360  
 gctccggcca gtgcggcccg gcgcgcctgc tgcccaacgc catcgccgc ggcaagtggg 420  
 ggcgacctag tgggcccgcac ttccgctgca tccccgaccg ctaccgcgcg cagcgcgtgc 480  
 agctgctgtg tcccgggtgt gaggcgcgc gcgcgcgcaa ggtgcgcctg gtggcctcgt 540  
 gcaagtgcaa gcgcctcacc cgcttcaca accagtcgga gctcaaggac ttcgggaccg 600  
 35 aggcgcgtcg gccgcagaag ggccggaagc gcgcgccccg cgcgcggagc gccaaagcca 660  
 accaggccga gctggagAAC gcctactaga gcccgccccg gccctcccc accggcgggc 720

gccccggccc tgaacccgcg cccacatctt ctgtcctctg cgcgtgggtt gattgtttat 780  
 atttcattgt aaatgcctgc aaccagggc agggggctga gacctccag gccctgagga 840  
 atccccggcg ccggcaaggc cccctcagc ccgccagctg aggggtccca cggggcaggg 900  
 gaggggaattg agagtcacag aactgagcc acgcagcccc gcctctgggg ccgcctacct 960  
 5 ttgctggtcc cacttcagag gaggcagaaa tggaagcatt ttcaccgccc tggggtttta 1020  
 agggagcggg gtgggagtg gaaagtccag ggactgggta agaaagttgg ataagattcc 1080  
 cccttgacc tcgctgccc tcagaaagcc tgaggcgtgc ccagagcaca agactggggg 1140  
 caactgtaga tgtggtttct agtctggct ctgccactaa cttgctgtgt aaccttgaac 1200  
 tacacaattc tccttcggga cctcaatttc cactttgtaa aatgaggggt gaggtgggaa 1260  
 10 taggatctcg aggagactat tggcatatga ttccaaggac tccagtgcct tttgaatggg 1320  
 cagaggtgag agagagagag agaaagagag agaataatg cagttgcatt gattcagtgc 1380  
 caaggtcact tccagaattc agagtgtgta tgctctcttc tgacagccaa agatgaaaaa 1440  
 caaacagaaa aaaaaagta aagagtctat ttatggctga catatttacg gctgacaaac 1500  
 tcctggaaga agctatgctg ctcccagcc tggcttcccc ggatgtttgg ctacctccac 1560  
 15 ccctccatct caaagaaata acatcatcca ttggggtaga aaaggagagg gtccgagggg 1620  
 ggtgggaggg atagaaatca catccgcccc aacttcccaa agagcagcat cctcccccg 1680  
 acccatagcc atgttttaaa gtcaccttc gaagagaagt gaaaggttca aggacactgg 1740  
 ccttgaggc ccgaggagc agccatcaca aactcacaga ccagcacatc ctttttgaga 1800  
 caccgccttc tgcccaccac tcacggacac atttctgcct agaaaacagc ttcttactgc 1860  
 20 tcttacatgt gatggcatat cttacactaa aagaatatta ttgggggaaa aactacaagt 1920  
 gctgtacata tgctgagaaa ctgcagagca taatagctgc caccacaaaa tctttttgaa 1980  
 aatcatttcc agacaacctc ttactttctg tgtagttttt aattgttaaa aaaaaaagt 2040  
 tttaaacaga agcacatgac atatgaaagc ctgcaggact ggtcgttttt ttggcaattc 2100  
 ttccacgtgg gacttgctca caagaatgaa agtagtggtt tttaaagagt taagttacat 2160  
 25 atttattttc tcacttaagt tatttatgca aaagtttttc ttgtagagaa tgacaatggt 2220  
 aatattgctt tatgaattaa cagtctgttc ttccagagtc cagagacatt gttaataaag 2280  
 acaatgaatc atgaccgaaa g 2301

&lt;210&gt; 6

30 &lt;211&gt; 213

&lt;212&gt; PRT

&lt;213&gt; Homo sapien

&lt;400&gt; 6

35 Met Gln Leu Pro Leu Ala Leu Cys Leu Ile Cys Leu Leu Val His Thr

1

5

10

15

Ala Phe Arg Val Val Glu Gly Gln Gly Trp Gln Ala Phe Lys Asn Asp  
                   20                                  25                                  30  
 Ala Thr Glu Ile Ile Arg Glu Leu Gly Glu Tyr Pro Glu Pro Pro Pro  
                   35                                  40                                  45  
 5 Glu Leu Glu Asn Asn Lys Thr Met Asn Arg Ala Glu Asn Gly Gly Arg  
                   50                                  55                                  60  
 Pro Pro His His Pro Phe Glu Thr Lys Asp Val Ser Glu Tyr Ser Cys  
                   65                                  70                                  75                                  80  
 Arg Glu Leu His Phe Thr Arg Tyr Val Thr Asp Gly Pro Cys Arg Ser  
 10                                  85                                  90                                  95  
 Ala Lys Pro Val Thr Glu Leu Val Cys Ser Gly Gln Cys Gly Pro Ala  
                   100                                  105                                  110  
 Arg Leu Leu Pro Asn Ala Ile Gly Arg Gly Lys Trp Trp Arg Pro Ser  
                   115                                  120                                  125  
 15 Gly Pro Asp Phe Arg Cys Ile Pro Asp Arg Tyr Arg Ala Gln Arg Val  
                   130                                  135                                  140  
 Gln Leu Leu Cys Pro Gly Gly Glu Ala Pro Arg Ala Arg Lys Val Arg  
                   145                                  150                                  155                                  160  
 Leu Val Ala Ser Cys Lys Cys Lys Arg Leu Thr Arg Phe His Asn Gln  
 20                                  165                                  170                                  175  
 Ser Glu Leu Lys Asp Phe Gly Thr Glu Ala Ala Arg Pro Gln Lys Gly  
                   180                                  185                                  190  
 Arg Lys Pro Arg Pro Arg Ala Arg Ser Ala Lys Ala Asn Gln Ala Glu  
                   195                                  200                                  205  
 25 Leu Glu Asn Ala Tyr  
                   210

&lt;210&gt; 7

&lt;211&gt; 2301

30

&lt;212&gt; DNA

&lt;213&gt; Homo sapien

&lt;400&gt; 7

agagcctgtg ctactggaag gtggcgtgcc ctccctctggc tggtagcatg cagctcccac 60  
 35 tggccctgtg tctcgtctgc ctgctggtac acacagcctt ccgtgtagtg gagggccagg 120  
 ggtggcaggc gttcaagaat gatgccacgg aaatcatccg cgagctcgga gagtaccaccg 180

	agcctccacc	ggagctggag	aacaacaaga	ccatgaaccg	ggcggagAAC	ggagggcggc	240
	ctccccacca	cccctttgag	accaaagacg	tgtccgagta	cagctgccgc	gagctgcact	300
	tcacccgcta	cgtgaccgat	gggccgtgcc	gcagcgccaa	gccggtcacc	gagctggtgt	360
	gctccggcca	gtgcggcccc	gcgcgcctgc	tgcccaacgc	catcggccgc	ggcaagtggg	420
5	ggcgacctag	tgggccccgac	ttccgctgca	tccccgaccg	ctaccgcgcg	cagcgcgtgc	480
	agctgctgtg	tcccgggtgg	gaggcgccgc	gcgcgcgcaa	ggtgcgcctg	gtggcctcgt	540
	gcaagtgcaa	gcgcctcacc	cgcttccaca	accagtcgga	gctcaaggac	ttcgggaccg	600
	aggccgctcg	gccgcagaag	ggccggaagc	cgcgcccccg	cgccccggagc	gccaaagcca	660
	accaggccga	gctggagAAC	gcctactaga	gccccgccgc	gccccctccc	accggcgggc	720
10	gccccggccc	tgaacccgcg	ccccacattt	ctgtcctctg	cgcggtggttt	gattgtttat	780
	atttcattgt	aaatgcctgc	aaccagggc	agggggctga	gaccttccag	gccctgagga	840
	atccccggcg	ccggcaaggc	ccccctcagc	ccgccagctg	aggggtcccc	cggggcaggg	900
	gagggAattg	agagtcacag	acactgagcc	acgcagcccc	gcctctgggg	ccgcctacct	960
	ttgctggtcc	cacttcagag	gaggcagaaa	tggaagcatt	ttcaccgccc	tggggtttta	1020
15	agggagcggg	gtgggagtgg	gaaagtccag	ggactggtta	agaaagtggg	ataagattcc	1080
	cccttgcacc	tcgctgcccc	tcagaaagcc	tgaggcgtgc	ccagagcaca	agactggggg	1140
	caactgtaga	tgtggtttct	agtcctggct	ctgccactaa	cttgctgtgt	aaccttgaac	1200
	tacacaattc	tccttcggga	cctcaatttc	cactttgtaa	aatgaggggtg	gaggtgggaa	1260
	taggatctcg	aggagactat	tggcatatga	ttccaaggac	tccagtgcct	tttgaatggg	1320
20	cagaggtgag	agagagagag	agaaagagag	agaatgaatg	cagttgcatt	gattcagtgc	1380
	caaggtcact	tccagaattc	agagttgtga	tgtctcttct	tgacagccaa	agatgaaaaa	1440
	caaacagaaa	aaaaaaagta	aagagtctat	ttatggctga	catattttacg	gctgacaaac	1500
	tcctggaaga	agctatgctg	cttcccagcc	tggcttcccc	ggatgttttg	ctacctccac	1560
	ccctccatct	caaagaaata	acatcatcca	ttggggtaga	aaaggagagg	gtccgagggt	1620
25	gggtgggagg	atagaaatca	catccgcccc	aacttcccaa	agagcagcat	ccctcccccg	1680
	acccatagcc	atgtttttaa	gtcaccttcc	gaagagaagt	gaaaggttca	aggacactgg	1740
	ccttgcaggc	ccgaggggagc	agccatcaca	aactcacaga	ccagcacatc	ccttttgaga	1800
	caccgccttc	tgcccaccac	tcacggacac	atttctgcct	agaaaacagc	ttcttactgc	1860
	tcttacatgt	gatggcatat	cttacactaa	agaatatta	ttgggggaaa	aactacaagt	1920
30	gctgtacata	tgttgagaaa	ctgcagagca	taatagctgc	cacccaaaaa	tctttttgaa	1980
	aatcatttcc	agacaacctc	ttactttctg	tgtagttttt	aattgtttaa	aaaaaaaaagt	2040
	tttaaacaga	agcacatgac	atatgaaagc	ctgcaggact	ggtcgttttt	ttggcaattc	2100
	ttccacgtgg	gacttgtcca	caagaatgaa	agtagtgggt	tttaaagagt	taagttacat	2160
	atatttttcc	tcacttaagt	tatttatgca	aaagtttttc	ttgtagagaa	tgacaatgtt	2220
35	aatattgctt	tatgaattaa	cagtctgttc	ttccagagtc	cagagacatt	gttaataaag	2280
	acaatgaatc	atgaccgaaa	g				2301

&lt;210&gt; 8

&lt;211&gt; 213

&lt;212&gt; PRT

5 &lt;213&gt; Homo sapien

&lt;400&gt; 8

Met Gln Leu Pro Leu Ala Leu Cys Leu Val Cys Leu Leu Val His Thr  
 1 5 10 15  
 10 Ala Phe Arg Val Val Glu Gly Gln Gly Trp Gln Ala Phe Lys Asn Asp  
 20 25 30  
 Ala Thr Glu Ile Ile Arg Glu Leu Gly Glu Tyr Pro Glu Pro Pro Pro  
 35 40 45  
 Glu Leu Glu Asn Asn Lys Thr Met Asn Arg Ala Glu Asn Gly Gly Arg  
 15 50 55 60  
 Pro Pro His His Pro Phe Glu Thr Lys Asp Val Ser Glu Tyr Ser Cys  
 65 70 75 80  
 Arg Glu Leu His Phe Thr Arg Tyr Val Thr Asp Gly Pro Cys Arg Ser  
 85 90 95  
 20 Ala Lys Pro Val Thr Glu Leu Val Cys Ser Gly Gln Cys Gly Pro Ala  
 100 105 110  
 Arg Leu Leu Pro Asn Ala Ile Gly Arg Gly Lys Trp Trp Arg Pro Ser  
 115 120 125  
 Gly Pro Asp Phe Arg Cys Ile Pro Asp Arg Tyr Arg Ala Gln Arg Val  
 25 130 135 140  
 Gln Leu Leu Cys Pro Gly Gly Glu Ala Pro Arg Ala Arg Lys Val Arg  
 145 150 155 160  
 Leu Val Ala Ser Cys Lys Cys Lys Arg Leu Thr Arg Phe His Asn Gln  
 165 170 175  
 30 Ser Glu Leu Lys Asp Phe Gly Thr Glu Ala Ala Arg Pro Gln Lys Gly  
 180 185 190  
 Arg Lys Pro Arg Pro Arg Ala Arg Ser Ala Lys Ala Asn Gln Ala Glu  
 195 200 205  
 Leu Glu Asn Ala Tyr  
 35 210

<210> 9  
 <211> 642  
 <212> DNA  
 <213> Cercopithecus pygerythrus

5

&lt;400&gt; 9

	atgcagctcc cactggccct gtgtcttgtc tgctgtgtgg tacacgcagc cttccgtgta	60
	gtggagggcc aggggtggca ggccttcaag aatgatgccca cggaaatcat ccccgagctc	120
	ggagagtacc ccgagcctcc accggagctg gagaacaaca agaccatgaa ccgggaggag	180
10	aatggagggc ggcctcccca ccaccccttt gagaccaaag acgtgtccga gtacagctgc	240
	cgagagctgc acttcacccg ctacgtgacc gatgggcccgt gccgcagcgc caagccagtc	300
	accgagttgg tgtgtctccgg ccagtgcggc ccggcacgcc tgctgcccga cgccatcggc	360
	cgcggaagt ggtggcgccc gagtgggccc gacttcgct gcacccccga ccgctaccgc	420
	gcgcagcgtg tgcagctgct gtgtcccggg ggtgccgcgc cgcgcgcgcg caaggtgcgc	480
15	ctgggtggcct cgtgcaagtg caagcgctc acccgcttcc acaaccagtc ggagctcaag	540
	gacttcggtc ccgaggccgc tcggccgcag aagggccgga agccgcggcc ccgcgcccgg	600
	ggggccaaag ccaatcaggc cgagctggag aacgcctact ag	642

<210> 10  
 20 <211> 213  
 <212> PRT  
 <213> Cercopithecus pygerythrus

&lt;400&gt; 10

25	Met Gln Leu Pro Leu Ala Leu Cys Leu Val Cys Leu Leu Val His Ala
	1 5 10 15
	Ala Phe Arg Val Val Glu Gly Gln Gly Trp Gln Ala Phe Lys Asn Asp
	20 25 30
	Ala Thr Glu Ile Ile Pro Glu Leu Gly Glu Tyr Pro Glu Pro Pro Pro
30	35 40 45
	Glu Leu Glu Asn Asn Lys Thr Met Asn Arg Ala Glu Asn Gly Gly Arg
	50 55 60
	Pro Pro His His Pro Phe Glu Thr Lys Asp Val Ser Glu Tyr Ser Cys
	65 70 75 80
35	Arg Glu Leu His Phe Thr Arg Tyr Val Thr Asp Gly Pro Cys Arg Ser
	85 90 95

Ala Lys Pro Val Thr Glu Leu Val Cys Ser Gly Gln Cys Gly Pro Ala  
 100 105 110  
 Arg Leu Leu Pro Asn Ala Ile Gly Arg Gly Lys Trp Trp Arg Pro Ser  
 115 120 125  
 5 Gly Pro Asp Phe Arg Cys Ile Pro Asp Arg Tyr Arg Ala Gln Arg Val  
 130 135 140  
 Gln Leu Leu Cys Pro Gly Gly Ala Ala Pro Arg Ala Arg Lys Val Arg  
 145 150 155 160  
 Leu Val Ala Ser Cys Lys Cys Lys Arg Leu Thr Arg Phe His Asn Gln  
 10 165 170 175  
 Ser Glu Leu Lys Asp Phe Gly Pro Glu Ala Ala Arg Pro Gln Lys Gly  
 180 185 190  
 Arg Lys Pro Arg Pro Arg Ala Arg Gly Ala Lys Ala Asn Gln Ala Glu  
 195 200 205  
 15 Leu Glu Asn Ala Tyr  
 210

&lt;210&gt; 11

&lt;211&gt; 638

20

&lt;212&gt; DNA

&lt;213&gt; Mus musculus

&lt;400&gt; 11

atgcagccct cactagcccc gtgcctcatc tgcctacttg tgcacgctgc cttctgtgct 60  
 25 gtggaggggcc aggggtggca agccttcagg aatgatgcca cagagggtcat cccagggcctt 120  
 ggagagtacc ccgagcctcc tcctgagaac aaccagacca tgaaccgggc ggagaatgga 180  
 ggcagacctc cccaccatcc ctatgacgcc aaaggtgtgt ccgagtacag ctgccgcgag 240  
 ctgcactaca cccgcttcct gacagacggc ccatgccgca gcgccaagcc ggtcaccgag 300  
 ttggtgtgct ccggccagtgc cgcccccgcg cggctgtctgc ccaacgccat cgggcgcgctg 360  
 30 aagtgggtggc gcccgaaagg accggatttc cgctgcatcc cggatcgcta ccgcgcgcgag 420  
 cgggtgcagc tgctgtgccc cgggggcgcg gcgccgcgct cgcgcaaggt gcgtctggtg 480  
 gcctcgtgca agtgcaagcg cctcaccgcg ttccacaacc agtcggagct caaggacttc 540  
 gggccggaga ccgcgcggcc gcagaagggt cgcaagccgc ggcccggcgc ccggggagcc 600  
 aaagccaacc aggcggagct ggagaacgcc tactagag 638  
 35

&lt;210&gt; 12



&lt;211&gt; 211

&lt;212&gt; PRT

&lt;213&gt; Mus musculus

5 &lt;400&gt; 12

Met Gln Pro Ser Leu Ala Pro Cys Leu Ile Cys Leu Leu Val His Ala  
 1 5 10 15  
 Ala Phe Cys Ala Val Glu Gly Gln Gly Trp Gln Ala Phe Arg Asn Asp  
 20 25 30  
 10 Ala Thr Glu Val Ile Pro Gly Leu Gly Glu Tyr Pro Glu Pro Pro Pro  
 35 40 45  
 Glu Asn Asn Gln Thr Met Asn Arg Ala Glu Asn Gly Gly Arg Pro Pro  
 50 55 60  
 His His Pro Tyr Asp Ala Lys Asp Val Ser Glu Tyr Ser Cys Arg Glu  
 15 65 70 75 80  
 Leu His Tyr Thr Arg Phe Leu Thr Asp Gly Pro Cys Arg Ser Ala Lys  
 85 90 95  
 Pro Val Thr Glu Leu Val Cys Ser Gly Gln Cys Gly Pro Ala Arg Leu  
 100 105 110  
 20 Leu Pro Asn Ala Ile Gly Arg Val Lys Trp Trp Arg Pro Asn Gly Pro  
 115 120 125  
 Asp Phe Arg Cys Ile Pro Asp Arg Tyr Arg Ala Gln Arg Val Gln Leu  
 130 135 140  
 Leu Cys Pro Gly Gly Ala Ala Pro Arg Ser Arg Lys Val Arg Leu Val  
 25 145 150 155 160  
 Ala Ser Cys Lys Cys Lys Arg Leu Thr Arg Phe His Asn Gln Ser Glu  
 165 170 175  
 Leu Lys Asp Phe Gly Pro Glu Thr Ala Arg Pro Gln Lys Gly Arg Lys  
 180 185 190  
 30 Pro Arg Pro Gly Ala Arg Gly Ala Lys Ala Asn Gln Ala Glu Leu Glu  
 195 200 205  
 Asn Ala Tyr  
 210

35 &lt;210&gt; 13

&lt;211&gt; 674

&lt;212&gt; DNA

&lt;213&gt; Rattus norvegicus

&lt;400&gt; 13

5 gaggaccgag tgcccttcct ccttctggca ccatgcagct ctactagcc ccttgccctg 60  
 cctgcctgct tgtacatgca gccttcgttg ctgtggagag ccaggggtgg caagccttca 120  
 agaatgatgc cacagaaatc atcccgggac tcagagagta ccagagacct cctcaggaac 180  
 tagagaacaa ccagaccatg aaccggggcg agaacggagg cagaccccc caccatcctt 240  
 atgacaccaa agacgtgtcc gagtacagct gccgcgagct gcactacacc cgcttcgtga 300  
 10 ccgacggccc gtgccgcagt gccaagccgg tcaccgagtt ggtgtgctcg ggccagtgcg 360  
 gccccgcgcg gctgctgccc aacgccatcg ggcgctgaa gtggtggcgc ccgaacggac 420  
 ccgacttccg ctgcatcccg gatcgctacc gcgcgcagcg ggtgcagctg ctgtgccccg 480  
 gcggcgcggc gccgcgctcg cgcaaggtgc gtctggtggc ctctgcaag tgcaagcgcc 540  
 tcaccgctt ccacaaccag tcggagctca aggacttcgg acctgagacc gcgcggccgc 600  
 15 agaagggctc caagccgcgg ccccgcgccc ggggagccaa agccaaccag gcggagctgg 660  
 agaacgccta ctag 674

&lt;210&gt; 14

&lt;211&gt; 213

20

&lt;212&gt; PRT

&lt;213&gt; Rattus norvegicus

&lt;400&gt; 14

Met Gln Leu Ser Leu Ala Pro Cys Leu Ala Cys Leu Leu Val His Ala  
 25 1 5 10 15  
 Ala Phe Val Ala Val Glu Ser Gln Gly Trp Gln Ala Phe Lys Asn Asp  
 20 25 30  
 Ala Thr Glu Ile Ile Pro Gly Leu Arg Glu Tyr Pro Glu Pro Pro Gln  
 35 40 45  
 30 Glu Leu Glu Asn Asn Gln Thr Met Asn Arg Ala Glu Asn Gly Gly Arg  
 50 55 60  
 Pro Pro His His Pro Tyr Asp Thr Lys Asp Val Ser Glu Tyr Ser Cys  
 65 70 75 80  
 Arg Glu Leu His Tyr Thr Arg Phe Val Thr Asp Gly Pro Cys Arg Ser  
 35 85 90 95  
 Ala Lys Pro Val Thr Glu Leu Val Cys Ser Gly Gln Cys Gly Pro Ala

100 105 110  
 Arg Leu Leu Pro Asn Ala Ile Gly Arg Val Lys Trp Trp Arg Pro Asn  
 115 120 125  
 Gly Pro Asp Phe Arg Cys Ile Pro Asp Arg Tyr Arg Ala Gln Arg Val  
 5 130 135 140  
 Gln Leu Leu Cys Pro Gly Gly Ala Ala Pro Arg Ser Arg Lys Val Arg  
 145 150 155 160  
 Leu Val Ala Ser Cys Lys Cys Lys Arg Leu Thr Arg Phe His Asn Gln  
 165 170 175  
 10 Ser Glu Leu Lys Asp Phe Gly Pro Glu Thr Ala Arg Pro Gln Lys Gly  
 180 185 190  
 Arg Lys Pro Arg Pro Arg Ala Arg Gly Ala Lys Ala Asn Gln Ala Glu  
 195 200 205  
 Leu Glu Asn Ala Tyr  
 15 210  
  
 <210> 15  
 <211> 532  
 <212> DNA  
 20 <213> Bos torus  
  
 <400> 15  
 agaatgatgc cacagaaatc atccccgagc tgggcgagta ccccgagcct ctgccagagc 60  
 tgaacaacaa gaccatgaac cgggcgaggaga acggagggag acctccccac cacccttttg 120  
 25 agaccaaaga cgctcccgag tacagctgcc gggagctgca cttcaccgcg tacgtgaccg 180  
 atggggccgtg ccgcagcgcc aagccggtca ccgagctggt gtgctcgggc cagtgcggcc 240  
 cggcgcgccct gctgccccaac gccatcggcc gcggcaagtg gtggcgccca agcgggccccg 300  
 acttccgctg catccccgac cgctaccgcg cgcagcgggt gcagctgttg tgtcctggcg 360  
 gcgcggcgcc gcgcgcgcgc aaggtgcgcc tgggtggcctc gtgcaagtgc aagcgccctca 420  
 30 ctgcgttcca caaccagtcc gagctcaagg acttcggggc cgaggccgcg cggccgcaaa 480  
 cgggcccggaa gctgcggccc cgcgcccggg gcaccaaagc cagccggggc ga 532  
  
 <210> 16  
 <211> 176  
 35 <212> PRT  
 <213> Bos torus

15

&lt;400&gt; 16

Asn Asp Ala Thr Glu Ile Ile Pro Glu Leu Gly Glu Tyr Pro Glu Pro  
 1 5 10 15  
 5 Leu Pro Glu Leu Asn Asn Lys Thr Met Asn Arg Ala Glu Asn Gly Gly  
 20 25 30  
 Arg Pro Pro His His Pro Phe Glu Thr Lys Asp Ala Ser Glu Tyr Ser  
 35 40 45  
 Cys Arg Glu Leu His Phe Thr Arg Tyr Val Thr Asp Gly Pro Cys Arg  
 10 50 55 60  
 Ser Ala Lys Pro Val Thr Glu Leu Val Cys Ser Gly Gln Cys Gly Pro  
 65 70 75 80  
 Ala Arg Leu Leu Pro Asn Ala Ile Gly Arg Gly Lys Trp Trp Arg Pro  
 85 90 95  
 15 Ser Gly Pro Asp Phe Arg Cys Ile Pro Asp Arg Tyr Arg Ala Gln Arg  
 100 105 110  
 Val Gln Leu Leu Cys Pro Gly Gly Ala Ala Pro Arg Ala Arg Lys Val  
 115 120 125  
 Arg Leu Val Ala Ser Cys Lys Cys Lys Arg Leu Thr Arg Phe His Asn  
 20 130 135 140  
 Gln Ser Glu Leu Lys Asp Phe Gly Pro Glu Ala Ala Arg Pro Gln Thr  
 145 150 155 160  
 Gly Arg Lys Leu Arg Pro Arg Ala Arg Gly Thr Lys Ala Ser Arg Ala  
 165 170 175  
 25

&lt;210&gt; 17

&lt;211&gt; 35828

&lt;212&gt; DNA

&lt;213&gt; Mus musculus

30

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;222&gt; (1)...(35828)

&lt;223&gt; n = A,T,C or G

35

&lt;400&gt; 17

	cgcgtttttg	tgagcagcaa	tattgcgctt	cgatgagcct	tggcgttgag	attgatacct	60
	ctgctgcaca	aaaggcaatc	gaccgagctg	gaccagcgca	ttcgtgacac	cgctccttc	120
	gaacttattc	gcaatggagt	gtcattcatc	aaggacngcc	tgatcgcaaa	tggtgctatc	180
	cacgcagcgg	caatcgaaaa	ccctcagccg	gtgaccaata	tctacaacat	cagccttggg	240
5	atcctgcgtg	atgagccagc	gcagaacaag	gtaaccgtca	gtgccgataa	gttcaaagtt	300
	aaacctgggtg	ttgataccaa	cattgaaacg	ttgatcgaaa	acgcgctgaa	aaacgctgct	360
	gaatgtgcgg	cgctggatgt	cacaaagcaa	atggcagcag	acaagaaagc	gatggatgaa	420
	ctggcttcct	atgtccgcac	ggccatcatg	atggaatgtt	cccccggtgg	tgttatctgg	480
	cagcagtgcc	gtcgatagta	tgcaattgat	aattattatc	atttgcggtt	cctttccggc	540
10	gatccgcctt	gttacggggc	ggcgacctcg	cggtttttcg	ctatttatga	aaattttccg	600
	gtttaaggcg	tttccgttct	tcttcgtcat	aacttaatgt	ttttatttaa	aataccctct	660
	gaaaagaaaag	gaaacgcacg	gtgctgaaaag	cgagcttttt	ggcctctgtc	gtttcctttc	720
	tctgtttttg	tccgtggaat	gaacaatgga	agtcaacaaa	aagcagagct	tatcgatgat	780
	aagcgggtcaa	acatgagaat	tcgcggccgc	ataatacgac	tcactatagg	gatcgacgcc	840
15	tactccccgc	gcatgaagcg	gaggagctgg	actccgcatg	cccagagacg	ccccccaacc	900
	cccaaagtgc	ctgacctcag	cctctaccag	ctctggcttg	ggcttggggc	gggtcaaggc	960
	taccacgttc	tcttaacagg	tggctgggct	gtctcttggc	cgcgcgctcat	gtgacagctg	1020
	cctagtctctg	cagtgaggtc	accgtggaat	gtctgccttc	gttgccatgg	caacgggatg	1080
	acgttacaat	ctgggtgtgg	agcttttctt	gtccgtgtca	ggaaatccaa	ataccctaaa	1140
20	ataccctaga	agaggaagta	gctgagccaa	ggctttcctg	gcttctccag	ataaagtttg	1200
	acttagatgg	aaaaaaacaa	aatgataaag	acccgagcca	tctgaaaatt	cctcctaatt	1260
	gcaccactag	gaaatgtgta	tattattgag	ctcgtatgtg	ttcttatttt	aaaaagaaaa	1320
	cttttagtcat	gttattaata	agaatttctc	agcagtggga	gagaaccaat	attaacacca	1380
	agataaaaagt	tggcatgata	cacattgcag	gaagatccac	gttgggtttt	catgaatgtg	1440
25	aagaccccat	ttattaaagt	cctaagctct	gtttttgcac	actaggaagc	gatggccggg	1500
	atggctgagg	ggctgtaagg	atctttcaat	gtcttacatg	tgtgtttcct	gtcctgcacc	1560
	taggacctgc	tgcctagcct	gcagcagagc	cagaggggtt	tcacatgatt	agtctcagac	1620
	acttgggggc	aggttgcatg	tactgcatcg	cttattttcca	tacggagcac	ctactatgtg	1680
	tcaaacacca	tatggtgttc	actcttcaga	acggtgggtg	tcatcatggt	gcatttgctg	1740
30	acggttggat	tgggtgtaga	gagctgagat	atatggacgc	actcttcagc	attctgtcaa	1800
	cgtggctgtg	cattcttgct	cctgagcaag	tggctaaaca	gactcacagg	gtcagcctcc	1860
	agctcagtcg	ctgcatagtc	ttagggaacc	tctcccagtc	ctccctacct	caactatcca	1920
	agaagccagg	gggcttggcg	gtctcaggag	cctgcttgct	gggggacagg	ttgttgagtt	1980
	ttatctgcag	taggttgctt	aggcatagtg	tcaggactga	tggctgcctt	ggagaacaca	2040
35	tcctttgccc	tctatgcaaa	tctgaccttg	acatggggggc	gctgctcagc	tgggaggatc	2100
	aactgcatac	ctaaagccaa	gcctaaagct	tcttcgtcca	cctgaaactc	ctggaccaag	2160

	gggcttccgg	cacatcctct	cagggcagtg	agggagctctg	tgtgagctgc	actttccaat	2220
	ctcagggcgt	gagaggcaga	gggaggtggg	ggcagagcct	tgcagctctt	tcctcccatc	2280
	tggacagcgc	tctggctcag	cagcccatat	gagcacaggc	acatccccac	cccacccccca	2340
	cctttcctgt	cctgcagaat	ttaggctctg	ttcacggggg	gggggggggg	ggggcagtc	2400
5	tatcctctct	taggtagaca	ggactctgca	ggagacactg	ctttgtaaga	tactgcagtt	2460
	taaatttgga	tgttgtaggg	ggaaagcgaa	gggcctcttt	gaccattcag	tcaaggtacc	2520
	ttctaactcc	catcgatttg	gggggctact	ctagtgtctag	acattgcaga	gagcctcaga	2580
	actgtagtta	ccagtgtggt	aggattgata	cttcagggag	cctgacatgt	gacagttcca	2640
	ttcttcaccc	agtcaccgaa	catttattca	gtacctaccc	cgtaacaggc	accgtagcag	2700
10	gtactgaggg	acggaccact	caaagaactg	acagaccgaa	gccttggaat	ataaacacca	2760
	aagcatcagg	ctctgccaac	agaacactct	ttaacactca	ggccctttta	cactcaggac	2820
	ccccaccccc	accccaagca	gttggcactg	ctatccacat	tttacagaga	ggaaaaacta	2880
	ggcacaggac	gatataagtg	gcttgcttaa	gcttgtctgc	atggtaaatg	gcagggctgg	2940
	attgagaccc	agacattcca	actctagggg	ctatttttct	tttttctcgt	tgttcgaatc	3000
15	tgggtcttac	tgggtaaact	caggctagcc	tcacactcat	atccttctcc	catggcttac	3060
	gagtgttagg	attccaggtg	tgtgctacca	tgtctgactc	cctgtagctt	gtctatacca	3120
	tcctcacaac	ataggaattg	tgatagcagc	acacacaccg	gaaggagctg	gggaaatccc	3180
	acagaggggt	ccgcaggatg	acaggcgaat	gcctacacag	aagggtggga	agggaaagcag	3240
	agggaaacagc	atgggcgtgg	gaccacaagt	ctatttgggg	aagctgccgg	taaccgtata	3300
20	tggctggggg	gaggggagag	gtcatgagat	gaggcaggaa	gagccacagc	aggcagcggg	3360
	tacgggctcc	ttattgccaa	gaggctcgga	tcttctcctc	cttctctcct	ccggggctgc	3420
	ctgttcattt	tccaccactg	cctcccatcc	aggtctgtgg	ctcaggacat	caccagctg	3480
	cagaaactgg	gcatacccca	cgtcctgaat	gctgccgagg	gcaggctcct	catgcacgtc	3540
	aacaccagtg	ctagcttcta	cgaggattct	ggcatcacct	acttgggcat	caaggccaat	3600
25	gatacgagg	agttcaacct	cagtgtttac	tttgaaagg	ccacagattt	cattgaccag	3660
	gcgctggccc	ataaaaaatg	taaggaaagt	acattccggc	acccatggag	cgtaagccct	3720
	ctgggacctg	cttctcccaa	agaggccccc	acttgaaaaa	ggttccagaa	agatcccaaa	3780
	atatgccacc	aactagggat	taagtgtcct	acatgtgagc	cgatgggggg	cactgcatat	3840
	agtctgtgcc	atagacatga	caatggataa	taatatttca	gacagagagc	aggagttagg	3900
30	tagctgtgct	cctttccctt	taattgagtg	tgcccatttt	tttattcatg	tatgtgtata	3960
	catgtgtgtg	cacacatgcc	ataggttgat	actgaacacc	gtcttcaatc	gttccccacc	4020
	ccaccttatt	ttttgaggca	gggtctcttc	cctgactctg	gggctcattg	gtttatctag	4080
	gctgctggcc	agtgagctct	ggagtctctg	ttttctctac	ctccctagcc	ctgggactgc	4140
	aggggcatgt	gctggggccag	gctttttatg	cgcgttgggg	atctgaactt	aggtccctag	4200
35	gcctgagcac	cgtaaagact	ctgccacatc	cccagcctgt	ttgagcaagt	gaaccattcc	4260
	ccagaattcc	cccagtgggg	ctttcctacc	cttttattgg	ctaggcattc	atgagtgggc	4320

	acctcgccag	aggaatgagt	ggccacgact	ggctcaggg	cagcagccta	gagatactgg	4380
	gttaagtctt	cctgccgctc	gctccctgca	gccgcagaca	gaaagtagga	ctgaatgaga	4440
	gctggctagt	ggtcagacag	gacagaaggc	tgagaggggc	acagggcaga	tgtcagcaga	4500
	gcagacaggt	tctccctctg	tgggggaggg	gtggcccaact	gcaggtgtaa	ttggccttct	4560
5	ttgtgctcca	tagaggcttc	ctgggtacac	agcagcttcc	ctgtcctggg	gattcccaaa	4620
	gagaactccc	taccactgga	cttacagaag	ttctattgac	tgggtgtaacg	gttcaacagc	4680
	tttggctctt	gggtggacgg	gcatactgct	gtatcagctc	aagagctcat	tcacgaatga	4740
	acacacacac	acacacacac	acacacacac	acacaagcta	attttgatat	gccttaacta	4800
	gctcagtga	tgggcatttc	tgaacatccc	tgaagttagc	acacatttcc	ctctgggtgt	4860
10	cctggcttaa	caccttctaa	atctatat	tatctttgct	gccctgttac	cttctgagaa	4920
	gcccctaggg	ccacttccct	tcgcacctac	attgctggat	ggtttctctc	ctgcagctct	4980
	taaatctgat	ccctctgcct	ctgagccatg	ggaacagccc	aataactgag	ttagacataa	5040
	aaacgtctct	agccaaaact	tcagctaaat	ttagacaata	aatcttactg	gttgtggaat	5100
	ccttaagatt	cttcatgacc	tccttcacat	ggcacgagta	tgaagcttta	ttacaattgt	5160
15	ttattgatca	aactaactca	taaaaagcca	gttgtctttc	acctgctcaa	ggaaggaaca	5220
	aaattcatcc	ttaactgatc	tgtgcacctt	gcacaatcca	tacgaatata	ttaagagtac	5280
	taagattttg	gttgtgagag	tcacatgtta	cagaatgtac	agctttgaca	aggtgcatcc	5340
	ttgggatgcc	gaagtgaact	gctgttccag	ccccctacct	tctgaggctg	ttttggaagc	5400
	aatgctctgg	aagcaacttt	aggaggtagg	atgctggaac	agcgggtcac	ttcagcatcc	5460
20	cgatgacgaa	tcccgctcaa	gctgtacatt	ctgtaacaga	ctgggaaagc	tgcagacttt	5520
	aaggccaggg	ccctatgggc	cctcttaata	cctgtcacac	ccaacccgag	cccttctcct	5580
	ccagccgttc	tgtgcttctc	actctggata	gatggagaac	acggccttgc	tagttaaagg	5640
	agtgaggctt	cacccttctc	acatggcagt	ggttggtcat	cctcattcag	ggaactctgg	5700
	ggcattctgc	ctttacttcc	tctttttgga	ctacagggaa	tatatgctga	cttgttttga	5760
25	ccttggtgat	ggggagactg	gatctttggg	ctggaatgtt	tcctgctagt	ttttcccat	5820
	cctttggcaa	accctatcta	tatcttacca	ctaggcatag	tggccctcgt	tctggagcct	5880
	gccttcaggc	tggttctcgg	ggaccatgtc	cctggtttct	ccccagcata	tgggtgtcac	5940
	agtgttcact	gcgggtgggt	gctgaacaaa	gcggggattg	catcccagag	ctccggtgcc	6000
	ttgtgggtac	actgctaaga	taaaatggat	actggcctct	ctctgaccac	ttgcagagct	6060
30	ctgggtgcctt	gtgggtacac	tgctaagata	aaatggatac	tggcctctct	ctatccactt	6120
	gcaggactct	agggaaacagg	aatccattac	tgagaaaacc	aggggctagg	agcagggagg	6180
	tagctgggca	gctgaagtgc	ttggcgacta	accaatgaat	accagagttt	ggatctctag	6240
	aatactctta	aaatctgggt	gggcagagtg	gcctgcctgt	aatcccagaa	ctcgggaggg	6300
	ggagacaggg	aatcatcaga	gcaaactggc	taaccagaat	agcaaaacac	tgagctctgg	6360
35	gctctgtgag	agatcctgcc	ttaacatata	agagagagaa	taaaacattg	aagaagacag	6420
	tagatgccaa	ttttaagccc	ccacatgcac	atggacaagt	gtgcgtttga	acacacatat	6480

	gcactcatgt gaaccaggca tgcacactcg ggcttatcac acacataatt tgaaagagag	6540
	agtgagagag gagagtgcac attagagttc acaggaaagt gtgagtgagc acacccatgc	6600
	acacagacat gtgtgccagg gagtaggaaa ggagcctggg tttgtgtata agagggagcc	6660
	atcatgtgtt tctaaggagg gcgtgtgaag gaggcgttgt gtgggctggg actggagcat	6720
5	ggttgttaact gagcatgctc cctgtgggaa acaggagggg ggccaccctg cagaggggtcc	6780
	cactgtccag cgggatcagt aaaagccctt gctgagaact ttaggtaata gccagagaga	6840
	gaaaggtagg aaagtggggg gactcccatc tctgatgtag gaggatctgg gcaagtagag	6900
	gtgcgtttga ggtagaaaaga ggggtgcaga ggagatgctc ttaattctgg gtcagcagtt	6960
	tctttccaaa taatgcctgt gaggaggtgt aggtgggtggc cattcactca ctcagcagag	7020
10	ggatgatgat gcccgggtgga tgctggaaat ggccgagcat caaccctggc tctggaagaa	7080
	ctccatcttt cagaaggaga gtggatctgt gtatggccag cggggtcaca ggtgcttggg	7140
	gcccctgggg gactcctagc actgggtgat gtttatcgag tgctcttgtg tgccaggcac	7200
	tggcctgggg ctttgtttct gtctctgttt tgtttcgttt tttgagacag actcttgcta	7260
	tgtatccgtg tcaatcttgg aatctcactg catagcccag gctgcggaga gaggggaggg	7320
15	caataggcct tgtaagcaag ccacacttca gagactagac tccaccctgc gaatgatgac	7380
	aggtcagagc tgagttccgg aagatTTTTT ttccagctgc caggtggagt gtggagtggc	7440
	agctagcggc aagggttagag ggcgagctcc ctgtgcagga gaaatgcaag caagagatgg	7500
	caagccagtg agttaagcat tctgtgtggg gagcaggtgg atgaagagag aggtctgggt	7560
	ttcgcctctg gggggggggg gaggggtggg gatgaggtga gaggggggca gctccctgca	7620
20	gtgtgatgag attttttctg acagtgacct ttggcctctc cctccccac ttcccttctt	7680
	tcctttcttc ccaccattgc tttccttgtc cttgagaaat tctgagtttc cacttactg	7740
	gtgatgcaga cggaaacaga agccgtgtgt gtgtgtgtgt gtgtgtgtgt gtgtgtgtgt	7800
	gtgtgtgtgt ttgtgtgtat gtgtgtgtgt gtgtttgtgt gtatgtgtgt cagtgggaat	7860
	ggctcatagt ctgcaggaag gtgggcagga aggaataagc tgtaggctga ggcagtgtgg	7920
25	gatgcaggga gagaggagag gagggatacc agagaaggaa attaaggag ctacaagagg	7980
	gcattgttgg ggtgtgtgtg tgtgtgtgtt gtttatattt gtattggaaa tacattcttt	8040
	taaaaaatac ttatccattt atttattttt atgtgcacgt gtgtgtgcct gcatgagttc	8100
	atgtgtgcca cgtgtgtgcg ggaacccttg gaggccacaa gggggcatct gatccccctg	8160
	aactggagtt ggaggaggtt gtgagtcctc tgacatgttt gctgggaact gaaccccggt	8220
30	cctatgcaag agcaggaagt gcagttatct gctgagccat ctctccagtc ctgaaatcca	8280
	ttctcttaaa atacacgtgg cagagacatg atgggattta cgtatggatt taatgtggcg	8340
	gtcattaagt tccggcacag gcaagcacct gtaaagccat caccacaacc gcaacagtga	8400
	atgtgacct ccccccatg ttcttcatgt cccctgtccc ctccatcctc cattctcaag	8460
	cacctcttgc tctgcctctg tcgctggaga acagtgtgca tctgcacact cttatgtcag	8520
35	tgaagtcaca cagcctgcac cccttcttgg tctgagtatt tgggttctga ctctgctatc	8580
	acacactact gtactgcatt ctctcgctct ctttttttaa acatattttt atttgtttgt	8640



	gtgtatgcac	atgtgccaca	tgtgtacaga	tactatggag	gccagaagag	gccatggccg	8700
	tccctggagc	tggagttaca	ggcagcgtgt	gagctgcctg	gtgtgggtgc	tgggaaccaa	8760
	acttgaatct	aaagcaagca	cttttaactg	ctgaggcagc	tctcagtacc	cttcttcatt	8820
	tctccgctg	ggttccattg	tatggacaca	tgtagctaga	atatcttgct	tatctaatta	8880
5	tgtacattgt	tttgtgctaa	gagagagtaa	tgctctatag	cctgagctgg	cctcaacctt	8940
	gccatcctcc	tgcctcagcc	tcctcctcct	gagtgcctagg	atgacaggcg	agtggtaact	9000
	tacatggttt	catgttttgt	tcaagactga	aggataacat	tcatacagag	aaggtctggg	9060
	tcacaaagt	tgcagttcac	tgaatggcac	aacctgtgat	caagaaacaa	aactcagggg	9120
	ctggagagat	ggcactgact	gctcttccag	aggtccggag	ttcaattccc	agcaaccaca	9180
10	tggtggctca	cagccatcta	taacgagatc	tgacgccctc	ttctgggtgtg	tctgaagaca	9240
	gctacagtgt	actcacataa	aataaataaa	tctttaaaac	acacacacac	acacaattac	9300
	caccccagaa	agcccaactcc	atgttccctc	ccacgtctct	gcctacagta	ctcccaggtt	9360
	accactgttc	aggtcttctaa	caacctgggt	tacttggggc	tcttttctgc	tctgtggagc	9420
	cacacatttg	tgtgcctcat	acacgttctt	tctagtaagt	tgcataattac	tctgcgtttt	9480
15	tacatgtatt	tatttattgt	agttgtgtgt	gcgtgtgggc	ccatgcatgg	cacagtgtgt	9540
	ggggatgtca	gagtattgtg	aacaggggac	agttcttttc	ttcaatcatg	tgggttccag	9600
	aggttgaact	caggtcatca	tgtgtggcag	caaatgcctt	tacccactga	gacatctcca	9660
	tattcttttt	ttttccctg	aggtgggggc	ttgttccata	gccc aaactg	gctttgcact	9720
	tgcagttcaa	agtgactccc	tgtctccacc	tcttagagta	ttggaattac	gatgtgtact	9780
20	accacacctg	actggatcat	taattctttg	atgggggcgg	ggaagcgcac	atgctgcagg	9840
	tgaagggatg	actggactgg	acatgagcgt	ggaagccaga	gaacagcttc	agtctaattgc	9900
	tctcccaact	gagctatttc	ggtttgccag	agaacaactt	acagaaagtt	ctcagtgccca	9960
	tgtggattcg	gggttgagg	tcaactcatc	agcttgacat	tggctcctct	acccactgag	10020
	ccttctcact	actctctacc	tagatcatta	attctttttt	aaaaagactt	attagggggc	10080
25	tggagagatg	gctcagccgt	taagagcacc	gaatgccctt	ccagagggtcc	tgagttcaat	10140
	tcccagcatg	ccattgctgg	gcagtagggg	gcgcaggtgt	tcaacgtgag	tagctgttgc	10200
	cagttttccg	cgggtggagaa	cctcttgaca	ccctgctgtc	cctgggtcatt	ctgggtgggt	10260
	gcatgggtgat	atgcttggtg	tatggaagac	tttgactgtt	acagtgaagt	tgggcttcca	10320
	cagttaccac	gtctccctg	tttcttgacg	gccgggtgct	tgtccattgc	cgcgagggtc	10380
30	acagccgctc	cccaacgcta	gttatcgctt	acctcatgat	gcggcagaag	atggacgtca	10440
	agtctgctct	gagtactgtg	aggcagaatc	gtgagatcgg	ccccaacgat	ggcttcctgg	10500
	cccaactctg	ccagctcaat	gacagactag	ccaaggaggg	caaggtgaaa	ctctaggggtg	10560
	cccacagcct	cttttgacga	ggcttgactg	ggagggccct	ggcagccatg	tttaggaaac	10620
	acagtatacc	cactccctgc	accaccagac	acgtgccccac	atctgtccca	ctctgggtcct	10680
35	cggggggccac	tccaccctta	gggagcacat	gaagaagctc	cctaagaagt	tctgctcctt	10740
	agccatcctt	tcctgtaatt	tatgtctctc	cctgaggtga	ggttcaggtt	tatgtccctg	10800

	tctgtggcat agatacatct cagtgaccca ggggtgggagg gctatcaggg tgcattggccc	10860
	gggacacggg cactcttcat gacccctccc ccacctgggt tcttctgtg tggccagaa	10920
	ccacgagcct ggtaaaggaa ctatgcaaac acaggccctg acctcccat gtctgttctt	10980
	ggtcctcaca gcccagacag ccctgctgag gcagacgaat gacattaagt tctgaagcag	11040
5	agtggagata gattagtac tagatttcca aaaagaagga aaaaaaaggc tgcattttaa	11100
	aattatttcc ttagaattaa agatactaca taggggacct tgggtaagca aatccattt	11160
	tcccagaggc tatcttgatt ctttgggaatg tttaaagtgt gccttgccag agagcttacg	11220
	atctatatct gctgcttcag agccttccct gaggatggct ctgttcttct gcttgtaga	11280
	agagcgatgc cttgggcagg gtttccccct tttcagaata cagggtgtaa agtccagcct	11340
10	attacaaaca aacaaacaaa caaacaacaa aaggacctc atttgagaa ttgcaaggat	11400
	tttctcctga attatagtgt tggtaggttc aagtcacac gccagtgtg tgcctcctg	11460
	gttgctattc taagaataat taggaggagg aacctagcca attgcagctc atgtccgtgg	11520
	gtgtgtgcac ggggtgcata gttggaaggg gtgcctgtcc ccttggggac agaaggaaaa	11580
	tgaaaggccc ctctgctcac cctggccatt tacgggaggc tctgctggtt ccacgggtgc	11640
15	tgtgcaggat cctgaaactg actcgtgga cagaaacgag acttggcggc accatgagaa	11700
	tggagagaga gagagcaaag aaagaaacag cctttaaaag aactttctaa ggggtggttt	11760
	tgaacctgc tggacctgt atgtgtgcac atttgccaga gattgaacat aatcctcttg	11820
	ggacttcacg ttctcattat ttgtatgtct ccgggggtcac gcagagccgt cagccaccac	11880
	cccagcacc ggcacatagg cgtctcataa aagcccattt tatgagaacc agagctgttt	11940
20	gagtacccc tgtatagaga gagttgtgt cgtggggcac ccggatccca gcagcctggt	12000
	tgcctgcctg taggatgtct tacaggagtt tgcagagaaa ccttctcttg agggaaagaa	12060
	atatcagggg tttttgttga atatttcaa ttcagcttta agtgaagac tcagcagtgt	12120
	tcattggttaa ggtaaggaa atgccttttc cagagctgct gcaagaggca ggagaagcag	12180
	acctgtctta ggatgtcact cccagggtta agacctctga tcacagcagg agcagagctg	12240
25	tgcagcctgg atggctcattg tcccctattc tgtgtgacca cagcaacct ggtcacatag	12300
	ggctgggtcat cttttttttt tttttttttt ttttttttg gcccagaatg aagtgacct	12360
	agccaagttg tgtacctcag tcttttagttt ccaagcggct ctcttgctca atacaatgtg	12420
	cattttcaaaa taacactgta gagttgacag aactggttca tgtgttatga gagaggaaaa	12480
	gagaggaaa aacaaaacaa aacaaaacac cacaaacaa aaacatctgg gctagccagg	12540
30	catgattgca atgtctacag gccagttca tgagaggcag agacaggaag accgccgaaa	12600
	ggtcaaggat agcatggtct acgtatcgag actccagcca gggctacggt cccaagatcc	12660
	taggttttgg attttgggt tgggtttttg agacagggtt tctctgtgta gccctggctg	12720
	tcctggaact cgctctgtag accaggctgg cctcaaactt agagatctgc ctgactctgc	12780
	ctttgagggc tgggacgaat gccaccactg cccaactaag attccattaa aaaaaaaaaa	12840
35	agttcaagat aattaagagt tgccagctcg ttaaagctaa gtagaagcag tctcaggcct	12900
	gctgcttgag gctgttcttg gcttggacct gaaatctgcc cccaacagtg tccaagtga	12960

	catgactttg agccatctcc agagaaggaa gtgaaaattg tggctcccca gtcgattggg	13020
	acacagtctc tctttgtcta ggtaacacat ggtgacacat agcattgaac tctccactct	13080
	gaggggtgggt tcccccccc ctgcctcttc tgggttggtc accccatagg acagccacag	13140
	gacagtcact agcacctact ggaaacctct ttgtgggaac atgaagaaag agcctttggg	13200
5	agattcctgg ctttccatta gggctgaaag tacaacggtt cttgggtggc tttgcctcgt	13260
	gtttataaaa ctagctacta ttcttcaggt aaaataccga tgttggtggaa aagccaaccc	13320
	cgtggctgcc cgtgagtagg ggggtggggtt ggggaatcctg gatagtgttc tatccatgga	13380
	aagtgggtgga ataggaatta aggggtgttc ccccccccc aacctcttcc tcagacccag	13440
	ccactttcta tgacttataa acatccaggt aaaaattaca aacataaaaa tggtttctct	13500
10	tctcaatctt ctaaagtctg cctgcctttt ccaggggtag gtctgtttct ttgctgttct	13560
	attgtcttga gagcacagac taacacttac caaatgaggg aactcttggc ccatactaag	13620
	gctcttctgg gctccagcac tcttaagtta ttttaagaat tctcacttgg ccttttagcac	13680
	acccgccacc cccaagtggg tgtggataat gccatggcca gcaggggggca ctggtgagggc	13740
	gggtgccttt ccaccttaag ttgcttatag tatttaagat gctaaatgtt ttaatcaaga	13800
15	gaagcactga tcttataata cgaggataag agattttctc acaggaaatt gtctttttca	13860
	taattctttt acaggctttg tctgatcgt agcatagaga gaatagctgg atatttaact	13920
	tgtattccat tttcctctgc cagcgttagg ttaactccgt aaaaagtgat tcagtggacc	13980
	gaagaggctc agagggcagg ggatgggtggg gtgaggcaga gcaactgtcac ctgccaggca	14040
	tgggaggtcc tgccatccgg gaggaaaagg aaagttagc ctctagtcta ccaccagtgt	14100
20	taacgcactc taaagtgtga accaaaataa atgtcttaca ttacaaagac gtctgttttg	14160
	tgtttccttt tgtgtgtttg ggctttttat gtgtgcttta taactgctgt ggtgggtgctg	14220
	ttgttagttt tgaggtagga tctcaggctg gccttgaact tctgatcgcc tgcctctgcc	14280
	cctgccccctg cccctgtccc tgctccaag tgctaggact aaaagcacat gccaccacac	14340
	cagtacagca tttttctaac atttaaaaaat aatcacctag gggctggaga gaggggtcca	14400
25	gctaagagtg cacactgtct ttgggtagga cctgagttta gttcccagaa cctatactgg	14460
	gtggctccag gtccagagga tccaggacct ctggcctcca tgggcatctg ctcttagcac	14520
	atacccatat acagatacac acataaaaaat aaaatgaagc ctttaaaaaac ctctaaaaac	14580
	ctagcccttg gaggtacgac tctggaaagc tggcactatg tgtaagtcca tctcatggtg	14640
	ttctggctaa cgtaagactt acagagacag aaaagaactc aggggtgtgct ggggggttggg	14700
30	atggaggaag agggatgagt agggggagca cggggaactt gggcagtga aattctttgc	14760
	aggacactag aggaggataa ataccagtca ttgcacccac tactggacaa ctccaggga	14820
	ttatgctggg tgaaaagaga agggccaggt tattggctgc attggctgca tttgcgtaac	14880
	atttttttaa attgaaaaga aaaagatgta aatcaagggt agatgagtgg ttgctgtgag	14940
	ctgagagctg ggggtgagtga gacatgtgga caactccatc aaaaagcgac agaaagaacg	15000
35	ggctgtgggtg acagctacct ctaatctcca cctccgggag gtgatcaagg ttagccctca	15060
	gctagcctgt ggtgcatgag accctgtttc aaaaacttta ataaagaaat aatgaaaaaa	15120

	gacatcaggg	cagatccttg	gggccaaagg	cggacaggcg	agtctcgtgg	taagggtcgtg	15180
	tagaagcgga	tgcattgagca	cgtgcccgag	gcatcatgag	agagccctag	gtaagtaagg	15240
	atggatgtga	gtgtgtcggc	gtcggcgcac	tgcacgtcct	ggctgtggtg	ctggactggc	15300
	atctttggtg	agctgtggag	gggaaatggg	tagggagatc	ataaaatccc	tccgaattat	15360
5	ttcaagaact	gtctattaca	attatctcaa	aatattaaaa	aaaaagaaga	attaaaaaac	15420
	aaaaaaccta	tccagggtgtg	gtggtgtgca	cctatagcca	cgggcacttg	gaaagctgga	15480
	gcaagaggat	ggcgagtttg	aaggatatctg	gggctgtaca	gcaagaccgt	cgtcccaaaa	15540
	ccaaacaaaa	cagcaaacc	attatgtcac	acaagagtgt	ttatagttag	cggcctcgct	15600
	gagagcatgg	ggtgggggtg	ggggtggggg	acagaaatat	ctaaactgca	gtcaataggg	15660
10	atccactgag	accctggggc	ttgactgcag	cttaaccttg	ggaaatgata	agggttttgt	15720
	gttgagtaaa	agcatcgatt	actgacttaa	cctcaaataga	agaaaaagaa	aaaaagaaaa	15780
	caacaaaagc	caaaccaagg	ggctgggtgag	atggctcagt	gggtaagagc	accctgactgc	15840
	tcttccgaag	gtccagagtt	caaatacccag	caaccacatg	gtggctcaca	accatctgta	15900
	acgagatatg	atgccctctt	ctgggtgtgtc	tgaagacagc	tacagtgtac	ttacatataa	15960
15	taaataaatc	ttaaaaaaaa	aaaaaaaaaaa	aaaagccaaa	ccgagcaaac	caggccccc	16020
	aacagaaggc	aggcacgacg	gcaggcacca	cgagccatcc	tgtgaaaagg	cagggtacc	16080
	catgggccga	ggagggtcca	gagagatagg	ctggtaagct	cagtttctct	gtataccctt	16140
	tttcttgttg	acactacttc	aattacagat	aaaataacaa	ataaacaaaa	tctagagcct	16200
	ggccactctc	tgctcgcttg	atctttctctg	ttacgtccag	cagggtggcgg	aagtgttcca	16260
20	aggacagatc	gcatcattaa	ggtggccagc	ataatctccc	atcagcaggt	ggtgctgtga	16320
	gaaccattat	ggtgctcaca	gaatcccggg	cccaggagct	gccctctccc	aagtctggag	16380
	caataggaaa	gctttctggc	ccagacaggg	ttaacagtcc	acattccaga	gcaggggaaa	16440
	aggagactgg	aggtcacaga	caaaagggcc	agcttctaac	aacttcacag	ctctggtagg	16500
	agagatagat	cacccccaac	aatggccaca	gctgggtttg	tctgccccga	aggaaactga	16560
25	cttaggaagc	aggtatcaga	gtcccccttc	tgaggggact	tctgtctgcc	ttgtaaagct	16620
	gtcagagcag	ctgcattgat	gtgtgggtga	cagaagatga	aaaggaggac	ccaggcagat	16680
	cgccacagat	ggaccggcca	cttacaagtc	gaggcagggtg	gcagagcctt	gcagaagctc	16740
	tgcagggtgga	cgacactgat	tcattaccca	gttagcatatc	cacagcgggc	taggcggacc	16800
	acagcctcct	tcccagtcctt	cctccagggc	tggggagtc	tccaaccttc	tgtctcagtg	16860
30	cagcttccgc	cagccccctc	tccttttgca	cctcagggtg	gaaccctccc	tcctctcctt	16920
	ctccctgtgg	catggccctc	ctgctactgc	aggctgagca	ttggatttct	ttgtgcttag	16980
	atagacctga	gatggctttc	tgatttatat	atatatatcc	atcccttggg	tcttacatct	17040
	aggacccaga	gctgtttgtg	ataccataag	aggctgggga	gatgatatgg	taagagtgtc	17100
	tgctgtacaa	gcatgaagac	atgagtctga	atccccagca	accatgtgga	aaaataacct	17160
35	tctaacctca	gagttgaggg	gaaaggcagg	tggattctgg	gggcttactg	gccagctagc	17220
	cagcctaacc	taaatgtctc	agtcagagat	cctgtctcag	ggaataactt	gggagaatga	17280

	ctgagaaaaga	cacctcctca	ggtctcccat	gcacccacac	agacacacgg	ggggggggta	17340
	atgtaataag	ctaagaaata	atgagggaaa	tgattttttg	ctaagaaatg	aaattctgtg	17400
	ttggccgcaa	gaagcctggc	caggggaagga	actgcctttg	gcacaccagc	ctataagtca	17460
	ccatgagttc	cctggctaag	aatcacatgt	aatggagccc	aggteccctc	tgcctggtgg	17520
5	ttgcctctcc	cactggtttt	gaagagaaat	tcaagagaga	tctccttggt	cagaattgta	17580
	gggtgctgagc	aatgtggagc	tggggccaat	gggattcctt	taaaggcatc	cttcccaggg	17640
	ctgggtcata	cttcaatagt	aggggtgctt	cacagcaagc	gtgagaccct	aggttagagt	17700
	ccccagaatc	tgcccccaac	cccccaaaaa	ggcatccttc	tgcctctggg	tgggtggggg	17760
	gagcaaacac	ctttaactaa	gaccattagc	tggcaggggt	aacaaatgac	cttggctaga	17820
10	ggaatttggt	caagctggat	tccgccttct	gtagaagccc	cacttgtttc	ctttgttaag	17880
	ctggcccaca	gtttgttttg	agaatgcctg	aggggcccag	ggagccagac	aattaaaagc	17940
	caagctcatt	ttgatatctg	aaaaccacag	cctgactgcc	ctgcccgtgg	gaggtactgg	18000
	gagagctggc	tgtgtccctg	cctcaccaac	gccccccccc	ccaacacaca	ctcctcggtt	18060
	cacctgggag	gtgccagcag	caatttgga	gtttactgag	cttgagaagt	cttgggaggg	18120
15	ctgacgctaa	gcacacccct	tctccacccc	ccccacccc	accccctga	ggaggaggg	18180
	gaggaaacat	gggaccagcc	ctgctccagc	ccgtccttat	tggctggcat	gaggcagagg	18240
	gggctttaaa	aaggcaaccg	tatctaggct	ggacactgga	gcctgtgcta	ccgagtcccc	18300
	tctccacct	ggcagcatgc	agccctcact	agccccgtgc	ctcatctgcc	tacttgtgca	18360
	cgctgccttc	tgtgctgtgg	agggccaggg	gtggcaagcc	ttcaggaatg	atgccacaga	18420
20	ggtcatccca	gggcttgagg	agtaccccga	gcctcctcct	gagaacaacc	agaccatgaa	18480
	ccgggcgagg	aatggaggca	gacctcccca	ccatccctat	gacgccaaag	gtacgggatg	18540
	aagaagcaca	ttagtggggg	gggggggtcct	gggaggtgac	tgggggtggt	ttagcatctt	18600
	cttcagaggt	ttgtgtgggt	ggctagcctc	tgctacatca	gggcagggac	acatttgctt	18660
	ggaagaatac	tagcacagca	ttagaacctg	gagggcagca	ttgggggggct	ggtagagagc	18720
25	acccaaggca	gggtggaggc	tgaggctcagc	cgaagctggc	attaacacgg	gcatgggctt	18780
	gtatgatggt	ccagagaatc	tcttcctaag	gatgaggaca	caggctcagat	ctagctgctg	18840
	accagtgggg	aagtgatatg	gtgaggctgg	atgccagatg	ccatccatgg	ctgtactata	18900
	tcccacatga	ccaccacatg	aggtaaagaa	ggccccagct	tgaagatgga	gaaaccgaga	18960
	ggctcctgag	ataaagtcac	ctgggagtaa	gaagagctga	gactggaagc	tggtttgatc	19020
30	cagatgcaag	gcaaccctag	attgggtttg	ggtgggaacc	tgaagccagg	aggaatccct	19080
	ttagttcccc	cttgcccagg	gtctgctcaa	tgagcccaga	gggttagcat	taaaagaaca	19140
	gggtttgtag	gtggcatgtg	acatgagggg	cagctgagtg	aaatgtcccc	tgtatgagca	19200
	cagggtggcac	cacttgccct	gagcttgcac	cctgacccca	gctttgcctc	attcctgagg	19260
	acagcagaaa	ctgtggaggc	agagccagca	cagagagatg	cctgggggtgg	gggtgggggt	19320
35	atcacgcacg	gaactagcag	caatgaatgg	ggtgggggtgg	cagctggagg	gacactccag	19380
	agaaatgacc	ttgctgggtca	ccatttgtgt	gggaggagag	ctcattttcc	agcttgccac	19440

	cacatgctgt	ccctcctgtc	tcctagccag	taagggatgt	ggaggaaagg	gccaccccaa	19500
	aggagcatgc	aatgcagtca	cgtttttgca	gaggaagtgc	ttgacctaa	ggcactattc	19560
	ttggaaagcc	ccaaaactag	tccttccttg	ggcaaacagg	cctccccac	ataccacctc	19620
	tgcaggggtg	agtaaattaa	gccagccaca	gaaggggtgg	aaggcctaca	cctccccctt	19680
5	gttggtgccc	cccccccccc	gtgaaggtgc	atcctggcct	ctgccccctt	ggctttggta	19740
	ctgggatttt	ttttttcctt	ttatgtcata	ttgatcctga	caccatggaa	cttttggagg	19800
	tagacaggac	ccacacatgg	attagttaaa	agcctcccat	ccatctaagc	tcatggtagg	19860
	agatagagca	tgtccaagag	aggagggcag	gcacagacc	tagaagatat	ggctgggcat	19920
	ccaacccaat	ctccttcccc	ggagaacaga	ctctaagtca	gatccagcca	cccttgagta	19980
10	accagctcaa	ggtacacaga	acaagagagt	ctggtatata	gcaggtgcta	aacaaatgct	20040
	tgtggtagca	aaagctatag	gttttggtgc	agaactccga	ccaagtcgc	gagtgaagag	20100
	cgaaaggccc	tctactcgcc	accgccccgc	ccccacctgg	ggctctataa	cagatcactt	20160
	tcacccttgc	gggagccaga	gagccctggc	atcctaggta	gccccccccg	cccccccccc	20220
	gcaagcagcc	cagccctgcc	tttggggcaa	gttcttttct	cagcctggac	ctgtgataat	20280
15	gaggggggtg	gacgcgcgcg	ctttggtcgc	tttcaagtct	aatgaattct	tatccctacc	20340
	acctgccctt	ctaccccgtt	cctccacagc	agctgtcctg	atttattacc	ttcaattaac	20400
	ctccactcct	ttctccatct	cctgggatac	cgccccgtgc	ccagtggctg	gtaaaggagc	20460
	ttaggaagga	ccagagccag	gtgtggctag	aggctaccag	gcagggctgg	ggatgaggag	20520
	ctaaactgga	agagtgtttg	gttagtaggc	acaaagcctt	gggtgggata	cctagtaccg	20580
20	gagaagtgga	gatgggcgct	gagaagttca	agaccatcca	tccttaacta	cacagccagt	20640
	ttgaggccag	cctgggctac	ataaaaaccc	aatctcaaaa	gctgccaat	ctgattctgt	20700
	gccacgtagt	gcccgatgta	atagtggatg	aagtcgttga	atcctggggc	aacctatttt	20760
	acagatgtgg	ggaaaagcaa	ctttaagtac	cctgcccaca	gatcacaaag	aaagtaagtg	20820
	acagagctcc	agtgtttcat	ccctgggttc	caaggacagg	gagagagaag	ccaggggtgg	20880
25	atctcactgc	tccccgggtg	ctccttccta	taatccatac	agattcgaaa	gcgcagggca	20940
	ggtttgga	aaagagagaag	ggtggaagga	gcagaccagt	ctggcctagg	ctgcagcccc	21000
	tcacgcatcc	ctctctccgc	agatgtgtcc	gagtacagct	gccgcgagct	gcactacacc	21060
	cgcttcctga	cagacggccc	atgccgcagc	gccaaagccg	tcaccgagtt	ggtgtgctcc	21120
	ggccagtgcg	gccccgcgcg	gctgctgccc	aacgccatcg	ggcgcgtgaa	gtggtggcgc	21180
30	ccgaacggac	cggattttccg	ctgcatcccc	gatcgctacc	gcgcgcagcg	ggtgcagctg	21240
	ctgtgccccg	ggggcgcggc	gccgcgctcg	cgcaaggtgc	gtctgggtgg	ctcgtgcaag	21300
	tgcaagcgcc	tcaccgcgtt	ccacaaccag	tcggagctca	aggacttcgg	gccggagacc	21360
	gcgcggccgc	agaaggggtcg	caagccgcgg	cccggcgccc	ggggagccaa	agccaaccag	21420
	gcggagctgg	agaacgccta	ctagagcgag	cccgcgccta	tgcagcccc	gcgcgatccg	21480
35	attcgttttc	agtgtaaagc	ctgcagccca	ggccaggggt	gccaaacttt	ccagaccgtg	21540
	tggagttccc	agcccagtag	agaccgcagg	tccttctgcc	cgctgcgggg	gatggggagg	21600

	gggtggggtt cccgcggggc aggagaggaa gcttgagtcc cagactctgc ctagccccgg	21660
	gtgggatggg ggtcttttcta ccctcgccgg acctatacag gacaaggcag tgtttccacc	21720
	ttaaagggaa gggagtgtgg aacgaaagac ctgggactgg ttatggacgt acagtaagat	21780
	ctactccttc caccctaatg taaagcctgc gtgggctaga tagggtttct gaccctgacc	21840
5	tggccactga gtgtgatgtt gggctacgtg gttctctttt ggtacgggtct tctttgtaaa	21900
	atagggaccg gaactctgct gagattccaa ggattggggg acccgtgta gactggtgag	21960
	agagaggaga acaggggagg ggttagggga gagattgtgg tgggcaaccg cctagaagaa	22020
	gctgtttgtt ggctcccagc ctgcgccct cagaggtttg gcttccccca ctcttctctc	22080
	tcaaactctgc cttcaaatec atatctggga tagggaaggc cagggtccga gagatggtgg	22140
10	aagggccaga aatcacactc ctggccccc gaagagcagt gtcccgcccc caactgcctt	22200
	gtcatattgt aaagggattt tctacacaac agtttaaggt cggtggagga aactgggctt	22260
	gccagtcacc tcccatcctt gtcccttgcc aggacaccac ctctgcctg ccaccacg	22320
	acacatttct gtctagaaac agagcgtcgt cgtgctgtcc tctgagacag catatcttac	22380
	attaaaaaga ataatacggg gggggggggc ggagggcgca agtggtatac atatgctgag	22440
15	aagctgtcag gcgccacagc accaccaca atctttttgt aaatcatttc cagacacctc	22500
	ttactttctg tgtagatttt aattgttaaa aggggaggag agagagcgtt tgtaacagaa	22560
	gcacatggag gggggggtag ggggggtggg gctggtgagt ttggcgaact ttccatgtga	22620
	gactcatcca caaagactga aagccgcgtt ttttttttta agagttcagt gacatattta	22680
	ttttctcatt taagttattt atgccaacat ttttttcttg tagagaaagg cagtgttaat	22740
20	atcgctttgt gaagcacaag tgtgtgtggg tttttgtttt ttgttttttc cccgaccaga	22800
	ggcattgtta ataaagacaa tgaatctcga gcaggaggct gtggtcttgt tttgtcaacc	22860
	acacacaatg tctcgccact gtcactcac tcccttccct tggtcacaag acccaaacct	22920
	tgacaacacc tccgactgct ctctggtagc ccttgtggca atacgtgttt cctttgaaaa	22980
	gtcacattca tcttttcttt tgcaaacctg gctctcattc cccagctggg tcatcgcat	23040
25	acctcaccc cagcttccct ttagctgacc actctccaca ctgtcttcca aaagtgcacg	23100
	tttcaccgag ccagttccct ggtccaggtc atcccattgc tcttcttgc tccagacct	23160
	tctcccacaa agatgttcat ctcccactcc atcaagcccc agtggccctg cggctatccc	23220
	tgtctcttca gttagctgaa tctacttgct gacaccacat gaattccttc ccctgtctta	23280
	aggttcatgg aactcttgcc tgccccctgaa ccttcaggga ctgtcccagc gtctgatgtg	23340
30	tcctctctct tgtaaagccc caccctacta tttgattccc aattctagat ctcccttgt	23400
	tcattccttc acgggatagt gtctcatctg gccaaagtcc gcttgatatt gggataaatg	23460
	caaagccaag tacaattgag gaccagttca tcattgggccc aagctttttc aaaatgtgaa	23520
	ttttacacct atagaagtgt aaaagccttc caaagcagag gcaatgcctg gctcttccct	23580
	caacatcagg gctcctgctt tatgggtctg gtggggtagt acattcataa acccaacact	23640
35	aggggtgtga aagcaagatg attgggagtt cgaggccaat cttggctatg aggccctgtc	23700
	tcaacctctc ctccctccct ccagggtttt gttttgtttt gtttttttga tttgaaactg	23760

	caacacttta	aatccagtc	agtgcattct	tgcgtgaggg	gaactctatc	cctaataataa	23820
	gcttccatct	tgatttgtgt	atgtgcacac	tgggggttga	acctgggcct	ttgtacctgc	23880
	cgggcaagct	ctctactgct	ctaaacccag	ccctcactgg	ctttctgttt	caactcccaa	23940
	tgaattcccc	taaatgaatt	atcaatatca	tgtctttgaa	aaataccatt	gagtgtctgct	24000
5	ggtgtccctg	tggttccaga	ttccaggaag	gacttttcag	ggaatccagg	catcctgaag	24060
	aatgtcttag	agcaggaggc	catggagacc	ttggccagcc	ccacaaggca	gtgtggtgca	24120
	gaggggtgag	atggaggcag	gcttgcaatt	gaagctgaga	caggggtactc	aggattaaaa	24180
	agcttcccc	aaaacaattc	caagatcagt	tcctgggtact	tgcacctgtt	cagctatgca	24240
	gagcccagtg	ggcataggtg	aagacaccgg	ttgtactgtc	atgtactaac	tgtgcttcag	24300
10	agccggcgaga	gacaaataat	gttatgggtga	ccccagggga	cagtgattcc	agaaggaaca	24360
	cagaagagag	tgctgctaga	ggctgcctga	aggagaaggg	gtcccagact	ctctaagcaa	24420
	agactccact	cacataaaga	cacagggtga	gcagagctgg	ccgtggatgc	agggagccca	24480
	tccaccatcc	tttagcatgc	ccttgtattc	ccatcacatg	ccagggatga	ggggcatcag	24540
	agagtccaag	tgatgcccaa	acccaaacac	acctaggact	tgctttctgg	gacagacaga	24600
15	tgcaggagag	actaggttgg	gctgtgatcc	cattaccaca	aagagggaaa	aaacaaaaaa	24660
	caaacaaaca	aacaaaaaaa	aacaaaacaa	aacaaaaaaa	aaccaaggt	ccaaattgta	24720
	ggtcagggtta	gagtttattt	atggaaagtt	atattctacc	tccatgggggt	ctacaaggct	24780
	ggcgcccatc	agaaagaaca	aacaacaggc	tgatctggga	ggggtggtac	tctatggcag	24840
	ggagcacgtg	tgcttgggggt	acagccagac	acggggcctg	tattaatcac	agggcttgta	24900
20	ttaataggct	gagagtcaag	cagacagaga	gacagaagga	aacacacaca	cacacacaca	24960
	cacacacaca	cacacacaca	catgcacaca	ccactcactt	ctcactcgaa	gagcccctac	25020
	ttacattcta	agaacaaacc	attcctcctc	ataaaggaga	caaagttgca	gaaacccaaa	25080
	agagccacag	ggtccccact	ctctttgaaa	tgacttggac	ttgttgcagg	gaagacagag	25140
	gggtctgcag	aggcttcctg	ggtgaccag	agccacagac	actgaaatct	ggtgctgaga	25200
25	cctgtataaa	ccctcttcca	caggttccct	gaaaggagcc	cacattcccc	aacctgtct	25260
	cctgaccact	gaggatgaga	gcacttgggc	cttccccatt	cttggagtgc	accctggttt	25320
	ccccatctga	gggcacatga	ggtctcaggt	cttgggaaag	ttccacaagt	attgaaagtg	25380
	ttcttgtttt	gtttgtgatt	taatttaggt	gtatgagtgc	ttttgcttga	atatatgcct	25440
	gtgtagcatt	tacaagcctg	gtgcctgagg	agatcagaag	atggcatcag	ataccctgga	25500
30	actggacttg	cagacagtta	tgagccactg	tgtgggtgct	aggaacagaa	cctggatcct	25560
	ccggaagagc	agacagccag	cgctcttagc	cactaagcca	tcactgaggt	tctttctgtg	25620
	gctaaagaga	caggagacaa	aggagagttt	cttttagtca	ataggaccat	gaatgttcct	25680
	cgtaacgtga	gactagggca	gggtgatccc	ccagtgcac	cgatggccct	gtgtagttat	25740
	tagcagctct	agtcttattc	cttaataagt	cccagtttgg	ggcaggagat	atgtattccc	25800
35	tgctttgaag	tggctgaggt	ccagttatct	acttccaagt	acttgtttct	ctttctggag	25860
	ttggggaagc	tcctgcctg	cctgtaaatg	tgtccattct	tcaaccttag	acaagatcac	25920



	tttccctgag	cagtcaggcc	agtccaaagc	ccttcaattt	agctttcata	aggaacaccc	25980
	cttttggttg	gtggaggtag	cacttgccct	gaatcccagc	attaagaagg	cagagacagt	26040
	cggatctctg	tgagttcaca	gccagcctgg	tctacggagt	gagttccaag	acagccaggc	26100
	ctacacagag	aaaccctgtc	tcgaaaaaaa	caaaaacaaa	agaaataaag	aaaaagaaaa	26160
5	caaaaaacgaa	caaacagaaa	aacaagccag	agtgtttgtc	cccgtatttt	attaatcata	26220
	tttttgctcc	tttgccattt	tagactaaaa	gactcgggaa	agcaggtctc	tctctgtttc	26280
	tcacccggac	acaccagaaa	ccagatgtat	ggaagatggc	taatgtgctg	cagttgcaca	26340
	tctggggctg	ggtggattgg	ttagatggca	tgggctgggt	gtggttacga	tgactgcagg	26400
	agcaaggagt	atgtggtgca	tagcaaacga	ggaagtttgc	acagaacaac	actgtgtgta	26460
10	ctgatgtgca	ggtatgggca	catgcaagca	gaagccaagg	gacagcctta	gggtagtgtt	26520
	tccacagacc	cctccccctt	tttaacatgg	gcatctctca	ttggcctgga	gcttgccaac	26580
	tgggctgggc	tggctagctt	gtaggtccca	gggatctgca	tatctctgcc	tccttagtgc	26640
	tgggattaca	gtcatatatg	agcacacctg	gcttttttat	gtgggttctg	ggctttgaac	26700
	ccagatctga	gtgcttgcaa	ggcaatcggt	tgaatgactg	cttcatctcc	ccagaccctg	26760
15	ggattctact	ttctattaaa	gtattttctat	taaatcaatg	agcccctgcc	cctgcactca	26820
	gcagttctta	ggcctgctga	gagtcaagtg	gggagtgaga	gcaagcctcg	agaccccatc	26880
	agcgaagcag	aggacaaaga	aatgaaaact	tgggattcga	ggctcgggat	atggagatac	26940
	agaaagggtc	agggaaggaa	atgaaccaga	tgaatagagg	caggaagggg	agggccctgc	27000
	atacatggaa	cctggtgtac	atgttatctg	catgggggtt	gcattgcaat	ggctcttcag	27060
20	caggttcacc	acactgggaa	acagaagcca	aaaagaagag	taggtggtgt	tggagtcaga	27120
	tactgtcagt	catgcctgaa	gaaatggaag	caattaacga	tgcgccgcaa	ttaggatatt	27180
	agctccctga	agaaaggcaa	gaagctgggc	tgtgggcact	gaagggagct	ttgaatgatg	27240
	tcacattctc	tgtatgccta	gcagggcagt	attggagact	gagacttgac	ttgtgtgtcc	27300
	atatgattcc	tccttttctt	acagtcactc	ggggctcctg	agcttcgtcc	ttgtccaaga	27360
25	acctggagct	ggcagtgggc	agctgcagtg	atagatgtct	gcaagaaaga	tctgaaaaga	27420
	gggaggaaga	tgaaggaccc	agaggaccac	cgacctctgc	tgcccgacaa	agctgcagga	27480
	ccagtctctc	ctacagatgg	gagacagagg	cgagagatga	atggtcaggg	gaggagtcat	27540
	agaaaggaga	gggtgaggca	gagaccaaag	gagggaaaca	cttgtgctct	acagctactg	27600
	actgagtacc	agctgcgtgg	cagacagcca	atgccaaggc	tcggctgatc	atggcacctc	27660
30	gtgggactcc	tagcccagtg	ctggcagagg	ggagtgtctg	atgggtgcatg	gtttggatat	27720
	gatctgaatg	tggtccagcc	ctagtttctt	tccagttgct	gggataaagc	accctgacca	27780
	aagctacttt	tttggttggt	tggttttggt	tggttttggt	tggtttttcg	aggcaggggt	27840
	tctctgtatc	accctagctg	tcctggaact	cactctgtag	accaggctgg	cctcgaactc	27900
	agaaatcccc	ctgcctctgc	ctcctaagtg	ctggaattaa	aggcctgcgc	caccactgcc	27960
35	ggcccaaagc	tactttaaga	gagagagagg	aatgtataag	tattataatt	ccaggttata	28020
	gttcattgct	gtagaattgg	agctttcata	ttccaggtaa	tctcccacag	acatgccaca	28080

	aaacaacctg	ttctacgaaa	tctctcatgg	actcccttcc	ccagtaattc	taaactgtgt	28140
	caaattctaca	agaaatagtg	acagtcacag	tctctaacgt	tttgggcatg	agtctgaagt	28200
	ctcattgcta	agtactggga	agatgaaaac	tttacctagt	gtcagcattt	ggagcagagc	28260
	ctttgggatt	tgagatgggc	ttttgcagag	ctcctaattg	ctacatggag	agagggggcc	28320
5	tgggagagac	ccatacacct	tttgctgcct	tatgtcacct	gacctgctcc	ttgggaagct	28380
	ctagcaagaa	ggccttccct	ggatcaccca	ccaccttgca	cctccagaac	tcagagccaa	28440
	attaaacttt	cttggttactg	tcgtcaaagc	acagtcgggc	tgggttgat	cactgtcaat	28500
	gggaaacaga	cttgccctgga	tggataactt	gtacattgca	taatgtctag	aaatgaaaag	28560
	tcctatagag	aaaaagaaaa	ttagctggca	cacagataga	ggccctggag	gaggctggct	28620
10	ttgtcctccc	cgaggaggtg	gcgagtaagg	tgtaaatgtt	catggatgta	aatgggcccc	28680
	tatatgaggg	tctggggtaa	caagaaggcc	tgtgaatata	aagcactgaa	ggtatgtcta	28740
	gtctggagaa	ggtcactaca	gagagtcttc	caactcagtg	cccatacaca	cacacacaca	28800
	cacacacaca	cacacacaca	cacacacaca	ccacaaagaa	aaaaaggaag	aaaaatctga	28860
	gagcaagtac	agtacttaaa	attgtgtgat	tgtgtgtgtg	actctgatgt	cacatgctca	28920
15	tcttgcccta	tgagttgaaa	accaaattggc	ccctgagagg	cataacaacc	acactgttgg	28980
	ctgtgtgctc	acgtttttct	taaagcgtct	gtctggtttg	ctgctagcat	caggcagact	29040
	tgcagcagac	tacatatgct	cagccctgaa	gtccttctag	ggtgcatgtc	tcttcagaat	29100
	ttcagaaagt	catctgtggc	tccaggaccg	cctgcactct	ccctctgccg	cgaggctgca	29160
	gactctaggc	tgggggtggaa	gcaacgctta	cctctgggac	aagtataaca	tgttggcttt	29220
20	tctttccctc	tgtggctcca	acctggacat	aaaatagatg	caagctgtgt	aataaatatt	29280
	tcctcccgtc	cacttagttc	tcaacaataa	ctactctgag	agcacttatt	aataggtggc	29340
	ttagacataa	gctttggctc	attccccac	tagctcttac	ttctttaact	ctttcaaacc	29400
	attctgtgtc	ttccacatgg	ttagttacct	ctccttccat	cctgggtcgc	ttcttccctc	29460
	gagtcgccct	cagtgtctct	aggtgatgct	tgtaaagatat	tctttctaca	aagctgagag	29520
25	tgggtggcact	ctgggagttc	aaagccagcc	tgatctacac	agcaagctcc	aggatatcca	29580
	gggcaatgtt	gggaaaacct	ttctcaaaca	aaaagagggg	ttcagttgtc	aggaggagac	29640
	ccatgggtta	agaagtctag	acgagccatg	gtgatgcata	cctttcatcc	aagcacttag	29700
	gaggcaaaga	aaggtgaaac	tctttgactt	tgaggccagc	taggttacat	agtgataccc	29760
	tgcttagtgt	gtgtgtgtgt	gtgtgtgtgt	gtgtgtgtgt	gtgtgttaatt	taaaagtcta	29820
30	aaaatgcatt	cttttaaaaa	tatgtataag	tatttgccctg	cacatatgta	tgtatgtatg	29880
	tataccatgt	gtgtgtctgg	tgctgaagga	ctaggcatag	actccctaga	actagagtca	29940
	tagacagttg	tgacactccc	caacccccca	ccatgtgggt	gcttgaagct	aaactcctgt	30000
	cctttgtaaa	gcagcaggtg	tctatgaacc	ctgaaccatc	tctccagtct	ccagatgtgc	30060
	attctcaaag	aggagtccct	catatttccc	taaactgaac	atccttatca	gtgagcatcc	30120
35	tcgagtcacc	aaagctactg	caaaccctct	tagggaacat	tcactattca	cttctacttg	30180
	gctcatgaaa	cttaagtaca	cacacacaaa	cacacacaca	cacacagagt	catgcactca	30240

	caaaagcatg catgtacacc attcttatta gactatgctt tgctaaaaga ctttcctaga	30300
	tactttaaaa catcacttct gccttttggg gggcagggtc caagattggg actggcgtag	30360
	tggaaactga acaaggtaga gatctagaaa tcacagcagg tcagaagggc cagcctgtac	30420
	aagagagagt tccacacctt ccaggaacac tgagcagggg gctgggacct tgccctctcag	30480
5	cccaagaaac tagtgcggtt cctgtatgca tgccctctcag agattccata agatctgcct	30540
	tctgccataa gatctcctgc atccagacaa gcctagggga agttgagagg ctgcctgagt	30600
	ctctcccaca ggcccccttct tgccctggcag tattttttta tctggaggag aggaatcagg	30660
	gtgggaatga tcaaatacaa ttatcaagga aaaagtaaaa aacatatata tatatatatt	30720
	aactgatcta gggagctggc tcagcagtta agagtctctg ctgcccttgc ttcagatctt	30780
10	gctttgattc ccagcaccca catgatggct ttcaactgta tctctgcttc caggggatcc	30840
	aacagcctct tctgacctcc atagacaaga cctagtcctc tgcaagagca ccaaagtctc	30900
	ttatctgttg atccatctct ctagcctcat gccagatcat ttaaaactac tggacactgt	30960
	cccattttac gaagatgtca ctgcccagtc atttgccatg agtggatatt tcgattcttt	31020
	ctatgttctc acccttgcaa ttataagaa agatatctgc atttgtctcc tgagagaaca	31080
15	aaggggtggg ggctactgag atggctctag gggtaaagg gcttgccaca aaatctgaca	31140
	acttaagttt ggtcttgga tccacatggt ggagagagag aagagattcc cgtaagtgtg	31200
	cctcaaactt cccacacatg tgctgtggct tatgtgtaac cccaataagt aaagatagtt	31260
	ttaaactact cataaggtag ggtttcttca tgaccccaag gaatgatgcc cctgatagag	31320
	cttatgctga aaccccatct ccattgtgcc atctggaaag agacaattgc atcccgga	31380
20	cagaatcttc atgaatggat taatgagcta ttaagaaagt ggcttggtta ttgcacatgc	31440
	tggcggcgta atgacctcca ccatgatgtt atccagcatg aaggctctca ccagaagtca	31500
	tacaaatctt cttaggcttc cagagtcgtg agcaaaaaaa gcacacctct aaataaatta	31560
	actagcctca ggtagttaac caccgaaaat gaaccaaggc agttctaata caaaaccact	31620
	tcccttccct gttcaaacca cagtgcctta ttatctaaaa gataaacttc aagccaagct	31680
25	tttaggttgc cagtatttat gtaacaacaa ggcccgttga cacacatctg taactcctag	31740
	tactgggcct caggggcaga gacagggtga gccctggagt ttgaattcca ggttctgtga	31800
	gaaactctgt ctgaaaagac aatatggtga gtgaccggg aggatatctg atattgactt	31860
	ctggccaaca cacagccatc tctgcacatc ttagattgca agccttttgc actaagtttg	31920
	gccagagtca gagtttgcaa gtgtttgtgg actgaatgca cgtgttgctg gtgatctaca	31980
30	aagtcaccct ccttctcaag ctagcagcac tggcttcggc cagctgctca ttcaagcctc	32040
	tttgcagagt catcacgggg atgggggagc agggccctc cctagaacac caagcctgtg	32100
	gttgtttatt caggacatta ttgaggcca agatgacaga taactctatc acttgccaa	32160
	cagtcgggtg ttgcggtgtt aggttatttc tgtgtctgca gaaaacagtg caacctggac	32220
	aaaagaaata aatgatatca tttttcatc aggcaactag attccgtggg acaaaaggct	32280
35	ccctggggaa cgaggccggg acagcgcggc tcctgagtcg ctatttccgt ctgtcaactt	32340
	ctctaacttc ttgatttctt cctctgtctt gtttcttcc tcttgctggg gccagtgga	32400

	gtctgtgtac	tcacagggag	gaggggtggca	aagccctggt	cctctacggg	ctggggggaag	32460
	ggggggaagct	gtcggccccag	tgactttttc	ccctttctct	ttttcttaga	aaccagtctc	32520
	aattttaagat	aatgagtctc	ctcattcacg	tgtgctcact	attcataggg	acttatccac	32580
	ccccgcccctg	tcaatctggc	taagtaagac	aagtcaaat	taaaagggaa	cgtttttcta	32640
5	aaaatgtggc	tggaaccgtgt	gccggcacga	aaccagggat	ggcgggtctaa	gttacatgct	32700
	ctctgccagc	cccgggtgcct	tttcccttcg	gaaaggagac	ccggaggtaa	aacgaagtgtg	32760
	ccaacttttg	atgatgggtgt	gcgccgggtg	actctttaa	atgtcatcca	tacctgggat	32820
	aggggaaggct	cttcagggag	tcatctagcc	ctcccttcag	gaaaagattc	cacttccggt	32880
	ttagtttagct	tccacctggt	cccttatccg	ctgtctctgc	ccactagtcc	tcatccatcc	32940
10	ggtttccgcc	ctcatccacc	ttgccctttt	agttcctaga	aagcagcacc	gtagtcttgg	33000
	caggtgggcc	attgggtcact	ccgctaccac	tgttaccatg	gccaccaagg	tgctatttaa	33060
	atatgagctc	actgagtcct	gcgggatggc	ttggttggtg	atatgcttgc	tgcaaaatcg	33120
	tgagaactgg	agttcaattc	ccagcacatg	gatgtatttc	cagcacctgg	aaggcagggg	33180
	gcagagatct	taaagctcct	ggccagacag	cccagcctaa	ttagtaatca	gtgagagacc	33240
15	ctgtctcaag	aaacaagatg	gaacatcaaa	ggtaaacctc	ttgtctccac	acacacaaat	33300
	acacacatgc	acatacatcc	acacacaggc	aaacacatgc	acacacctga	acaccctcca	33360
	caaatacata	cataaaaaaa	taaatacata	cacacatata	tacatacacc	aacattccct	33420
	ctccttagtc	tcttggttac	gctcttgtca	ccccactaa	ggcttcaact	tcttctatct	33480
	cttcatcttg	actcctctgt	actttgcatg	ccttttccag	caaaggcttt	tctttaaatc	33540
20	tccgtcatte	ataaactccc	tctaaatttc	ttcccctgcc	cttttctttc	tctctagggg	33600
	gataaagaca	cacactacaa	agtcaccgtg	ggaccagttt	attcacccac	ccaccctgc	33660
	ttctgttcat	ccggccagct	aagtagtcca	acctctctgg	tgtgttacc	tggaccctgg	33720
	cttcaccaca	gctcctccat	gctaccacgc	cctgcaaacc	ttcagcctag	cctctggttc	33780
	tccaaccagc	acaggccccag	tctggcttct	atgtcctaga	aatctccttc	attctctcca	33840
25	tttccctcct	gaatctacca	ccttctttct	cccttctcct	gacctcta	gtcttggtca	33900
	aacgattaca	aggaagccaa	tgaaattagc	agtttggggg	acctcagagt	cagcagggga	33960
	gctgggatga	attcacatct	ccaggccctt	gctttgctcc	ccggattctg	acaggcagtt	34020
	ccgaagctga	gtccaggaag	ctgaatttaa	aatcacactc	cagctggggt	ctgaggcagc	34080
	cctaccacat	cagctggccc	tgactgagct	gtgtctgggt	ggcagtgggtg	ctgggtgggtg	34140
30	tggtgggtgct	gggtgggtgg	gtgggtgggtg	tggtgggtgg	gggtgggtgg	tgtgtgtgtg	34200
	ttttctgctt	ttacaaaact	tttctaattc	ttatacaaag	gacaaatctg	cctcatatag	34260
	gcagaaagat	gacttatgcc	tatataagat	ataaagatga	ctttatgcca	cttatttagca	34320
	atagttactg	tcaaaagtaa	ttctatttat	acacccttat	acatgggtatt	gcttttgttg	34380
	gagactctaa	aatccagatt	atgtatttaa	aaaaaaatc	cccagtcctt	aaaagggtgaa	34440
35	gaatggaccc	agatagaagg	tcacggcaca	agtatggagt	cggagtgtgg	agtcctgcc	34500
	atgggtctgga	cagaagcatc	cagagagggg	ccaagacaaa	tgctctgcct	cctaaggaac	34560

	actggcagcc ctgatgaggt accagagatt gctaagtgga ggaatacagg atcagaccca	34620
	tggaggggct taaagcgtga ctgtagcagc cctccgctga ggggctccag gtgggcgccc	34680
	aagggtgctgc agtgggagcc acatgagagg tgatgtcttg gagtcacctc gggtagcatt	34740
	gttttagggag gtggggattt gtgggtgtgga gacaggcagc ctcaaggatg cttttcaaca	34800
5	atgggttgatg agttggaact aaaacagggg ccatcacact ggctcccata gctctgggct	34860
	tgccagcttc cacatctgcc cccaccccc tgtctggcac cagctcaagc tctgtgattc	34920
	tacacatcca aaagaggaag agtagcctac tgggcatgcc acctcttctg gaccatcagg	34980
	tgagagtgtg gcaagcccta ggctcctgtc caggatgcag ggctgccaga taggatgctc	35040
	agctatctcc tgagctggaa ctatttttagg aataaggatt atgcccgcgc ggggttggcc	35100
10	agcaccccag cagcctgtgc ttgctgtaaaa gcaagtgtctg ttgatttatc taaaaacaga	35160
	gccgtggacc caccacacagg acaagtatgt atgcatctgt ttcattgtatc tgaaaagcga	35220
	cacaaccatt ttccacatca tggcatcttc ctaaccccc tttttttttg ttttgttttt	35280
	ttgagacagg gtttctctgt gtagtcctgg ctgtcctgga actcactttg tagaccaggc	35340
	tggcctcgaa ctcagaaatc ctgggattaa aggtgtgtgc caccacgccc ggccctaacc	35400
15	cccattctta atgggtgatcc agtgggttgaa atttcggggc acacacatgt ccattaggga	35460
	ttagctgctg tcttctgagc tacctggtag aatctttatc ccctggggcc tgggctcctg	35520
	atccctgact cgggcccgat caagtccagt tcctggggcc gatcaagtcc agttcctggg	35580
	cccgaacaag tccagtccct agctcgatta gctcatcctg gctccctggc ctgttcttac	35640
	ttacactctt ccccttgctc tggacttggt gctttcttta ctcaagttgt ctgccacagt	35700
20	ccctaagcca cctctgtaag acaactaaga taatacttcc ctcaagcacg gaaagtcctg	35760
	agtcaccaca ccctctggag gtgtgtggac acatgttcat gcgtgtgggt gcgcttacgt	35820
	acgtgtgc	35828

&lt;210&gt; 18

25 &lt;211&gt; 9301

&lt;212&gt; DNA

&lt;213&gt; Homo sapien

&lt;400&gt; 18

30	tagaggagaa gtctttgggg agggtttgct ctgagcacac ccctttccct ccctccgggg	60
	ctgagggaaa catgggacca gccctgcccc agcctgtcct cattggctgg catgaagcag	120
	agaggggctt taaaaaggcg accgtgtctc ggctggagac cagagcctgt gctactggaa	180
	gggtggcgtgc ctcctctgg ctggtaccat gcagctccca ctggccctgt gtctcgctctg	240
	cctgctggta cacacagcct tccgtgtagt ggagggccag ggggtggcagg cgttcaagaa	300
35	tgatgccacg gaaatcatcc ccgagctcgg agagtacccc gagcctccac cggagctgga	360
	gaacaacaag accatgaacc gggcggagaa cggagggcgg cctccccacc acccctttga	420

	gaccaaaggt atgggggtgga ggagagaatt cttagtaaaa gatcctgggg aggttttaga	480
	aacttctctt tgggaggctt ggaagactgg ggtagaccca gtgaagattg ctggcctctg	540
	ccagcactgg tcgaggaaca gtcttgccctg gaggtggggg aagaatggct cgctggtgca	600
	gccttcaa at tcaggtgcag aggc atgagg caacagacgc tgggtgagagc ccagggcagg	660
5	gaggacgctg ggggtggtgag ggtatggcat caggggcatca gaacaggctc aggggctcag	720
	aaaagaaaag gtttcaaaga atctcctcct gggaaatatag gagccacgctc cagctgctgg	780
	taccactggg aaggggaaca ggtaagggag cctcccatcc acagaacagc acctgtgggg	840
	caccggacac tctatgctgg tgggtggctgt cccaccaca cagaccaca tcatggaatc	900
	cccaggaggt gaacccccag ctggaagggg aagaaacagg ttccaggcac tcagtaactt	960
10	ggtagtgaga agagctgagg tgtgaacctg gtttgatcca actgcaagat agccctgggtg	1020
	tgtggggggg tgtgggggac agatctccac aaagcagtgg ggaggaaggc cagagaggca	1080
	cccctgcagt gtgcattgcc catggcctgc ccaggagct ggcacttgaa ggaatgggag	1140
	ttttcggcac agtttttagcc cctgacatgg gtgcagctga gtccaggccc tggaggggag	1200
	agcagcatcc tctgtgcagg agtagggaca tctgtcctca gcagccacc cagtcccaac	1260
15	cttgccctcat tccaggggag ggagaaggaa gaggaacctt gggttcctgg tcaggcctgc	1320
	acagagaagc ccaggtgaca gtgtgcatct ggctctataa ttggcaggaa tcctgaggcc	1380
	atgggggctgt ctgaaatgac acttcagact aagagcttcc ctgtcctctg gccattatcc	1440
	aggtggcaga gaagtccact gccaggctc ctggacccca gccctccccg cctcacaacc	1500
	tgttgggact atgggggtgt aaaaagggca actgcatggg aggccagcca ggaccctccg	1560
20	tcttcaaaat ggaggacaag ggcgcctccc cccacagctc cccttctagg caaggtcagc	1620
	tgggctccag cgactgcctg aagggtgtga aggaacccaa acacaaaatg tccaccttgc	1680
	tggactccca cgagaggcca cagcccctga ggaagccaca tgctcaaaac aaagtcatga	1740
	tctgcagagg aagtgcctgg cctaggggag ctattctcga aaagccgcaa aatgccccct	1800
	tccctgggca aatgcccccc tgaccacaca cacattccag ccctgcagag gtgaggatgc	1860
25	aaaccagccc acagaccaga aagcagcccc agacgatggc agtggccaca tctcccctgc	1920
	tgtgcttgc ctccagagtg ggggtggggg gtggccttct ctgtccctc tctggtttgg	1980
	tcttaagact atttttcatt ctttcttgtc acattggaac tatccccatg aaacctttgg	2040
	gggtggactg gtactcacac gacgaccagc tatttaaaaa gctccaccc atctaagtcc	2100
	accataggag acatgggtcaa ggtgtgtgca ggggatcagg ccaggcctcg gagcccaatc	2160
30	tctgcctgcc caggaggtat caccatgagg cgcctattca gataacacag aacaagaaat	2220
	gtgcccagca gagagccagg tcaatgtttg tggcagctga acctgtaggt tttgggtcag	2280
	agctcagggc ccctatggta ggaaagtaac gacagtaaaa agcagccctc agctccatcc	2340
	cccagcccag cctcccatgg atgctcgaac gcagagcctc cactcttgcc ggagccaaaa	2400
	ggtgctggga cccaggggaa gtggagtcag gagatgcagc ccagcctttt gggcaagttc	2460
35	ttttctctgg ctgggcctca gtattctcat tgataatgag ggggttggac aactgcctt	2520
	tgattccttt caagtcta at gaattcctgt cctgatcacc tccccctcag tccctgcct	2580

	ccacagcagc	tgccctgatt	tattaccttc	aattaacctc	tactcctttc	tccatccccct	2640
	gtccacccccct	cccaagtggc	tggaaaagga	atttgggaga	agccagagcc	aggcagaagg	2700
	tgtgctgagt	acttaccctg	cccaggccag	ggaccctgcg	gcacaagtgt	ggcttaaatac	2760
	ataagaagac	cccagaagag	aaatgataat	aataatacat	aacagccgac	gctttcagct	2820
5	atatgtgcca	aatggtatct	tctgcattgc	gtgtgtaatg	gattaactcg	caatgcttgg	2880
	ggcggcccat	tttgagaca	ggaagaagag	agagggttaag	gaacttgccc	aagatgacac	2940
	ctgcagttag	cgatggagcc	ctggtgtttg	aacccagca	gtcatttggc	tccgagggga	3000
	caggggtg	aggagagctt	tccaccagct	ctagagcatc	tgggaccttc	ctgcaataga	3060
	tgttcagggg	caaaagcctc	tggagacagg	cttggcaaaa	gcagggctgg	ggtggagaga	3120
10	gacgggccc	tccagggcag	gggtggccag	gcgggcggcc	accctcacgc	gcgcctctct	3180
	ccacagacgt	gtccgagtag	agctgccg	agctgcactt	cacccgctac	gtgaccgatg	3240
	ggccgtgccg	cagcgccaag	ccggtcaccg	agctgggtgtg	ctccggccag	tgcgggccc	3300
	cgcgctgct	gcccacg	atcgccg	gcaagtgggtg	gcgacctagt	gggcccga	3360
	tccgctgcat	ccccgaccgc	taccgcgcgc	agcgctgca	gctgctgtgt	cccgggtgtg	3420
15	aggcgccg	cgcgcgcaag	gtgcgcctgg	tggcctcgtg	caagtgcaag	cgctcaccc	3480
	gcttcacaa	ccagtcggag	ctcaaggact	tcgggaccga	ggccgctcgg	ccgcagaagg	3540
	gccggaagcc	gcggccccgc	gcccggagcg	ccaaagccaa	ccaggccgag	ctggagaacg	3600
	cctactagag	cccgcccg	cccctcccca	ccggcggg	ccccggccct	gaacccgcgc	3660
	cccacatttc	tgctctctgc	gcgtgggtttg	attgtttata	tttcattgta	aatgcctgca	3720
20	acccagggca	gggggctgag	accttccagg	ccctgaggaa	tcccgggcgc	cggcaaggcc	3780
	cccctcagcc	cgccagctga	gggggtccac	ggggcagggg	aggggaattga	gagtcacaga	3840
	cactgagcca	cgcagccccg	cctctggggc	cgctacctt	tgctgggtccc	acttcagagg	3900
	aggcagaaat	ggaagcattt	tcaccgccct	ggggttttta	gggagcggtg	tgggagtggg	3960
	aaagtccagg	gactgggttaa	gaaagtgtga	taagattccc	ccttgcacct	cgctgccc	4020
25	cagaaagcct	gaggcgtagc	cagagcaca	gactgggggc	aactgtagat	gtgggtttcta	4080
	gtcctggctc	tgccactaac	ttgctgtgta	accttgaact	acacaattct	ccttcgggac	4140
	ctcaatttcc	actttgtaaa	atgaggggtg	aggtgggaat	aggatctcga	ggagactatt	4200
	ggcatatgat	tccaaggact	ccagtgcctt	ttgaatgggc	agaggtgaga	gagagagaga	4260
	gaaagagaga	gaatgaatgc	agttgcattg	attcagtgcc	aaggtcactt	ccagaattca	4320
30	gagttgtgat	gctctcttct	gacagccaaa	gatgaaaaac	aaacagaaaa	aaaaaagtaa	4380
	agagtctatt	tatggctgac	atatttacgg	ctgacaaact	cctggaagaa	gctatgctgc	4440
	ttcccagcct	ggcttcctcg	gatgtttggc	tacctccacc	cctccatctc	aaagaaataa	4500
	catcatccat	tggggtagaa	aaggagaggg	tccgaggggtg	gtgggagggga	tagaaatcac	4560
	atccgcccc	acttcccaaa	gagcagcatc	cctcccccca	cccatagcca	tgtttttaag	4620
35	tcaccttccg	aagagaagtg	aaagggttcaa	ggacactggc	cttgcaggcc	cgagggagca	4680
	gccatcaca	actcacagac	cagcacatcc	cttttgagac	accgccttct	gccaccact	4740

	cacggacaca	tttctgccta	gaaaacagct	tcttactgct	cttacatgtg	atggcatatc	4800
	ttacactaaa	agaatattat	tgggggaaaa	actacaagtg	ctgtacatat	gctgagaaac	4860
	tgcagagcat	aatagctgcc	acccaaaaat	ctttttgaaa	atcatttcca	gacaacctct	4920
	tactttctgt	gtagttttta	attgttaaaa	aaaaaaagtt	ttaaacagaa	gcacatgaca	4980
5	tatgaaagcc	tgcaggactg	gtcgtttttt	tggcaattct	tccacgtggg	acttgtccac	5040
	aagaatgaaa	gtagtggttt	ttaaagagtt	aagttacata	tttattttct	cacttaagtt	5100
	atztatgcaa	aagtttttct	tgtagagaat	gacaatgtta	atattgcttt	atgaattaac	5160
	agtctgttct	tccagagtcc	agagacattg	ttaataaaga	caatgaatca	tgaccgaaag	5220
	gatgtgggtc	cattttgtca	accacacatg	acgtcatttc	tgtcaaagtt	gacaccttc	5280
10	tcttggtcac	tagagctcca	accttgga	cacctttgac	tgctctctgg	tggcccttgt	5340
	ggcaattatg	tcttcctttg	aaaagtcag	tttatccctt	cctttccaaa	cccagaccgc	5400
	atctcttcac	ccagggcatg	gtaataacct	cagccttgta	tccttttagc	agcctccctt	5460
	ccatgctggc	ttccaaaatg	ctgttctcat	tgtatcactc	cctgctcaa	aagccttcca	5520
	tagctcccc	ttgcccagga	tcaagtgcag	tttccctatc	tgacatggga	ggccttctct	5580
15	gcttgactcc	cacctccac	tccaccaagc	ttcctactga	ctccaaatgg	tcatgcagat	5640
	cctgcttcc	ttagtttgcc	atccacactt	agcaccacca	ataactaatc	ctctttcttt	5700
	aggattcaca	ttacttgtca	tctcttcccc	taaccttcca	gagatgttcc	aatctcccat	5760
	gatccctctc	tcctctgagg	ttccagcccc	ttttgtctac	accactactt	tggttcctaa	5820
	ttctgttttc	catttgacag	tcattcatgg	aggaccagcc	tggccaagtc	ctgcttagta	5880
20	ctggcataga	caacacaaaag	ccaagtacaa	ttcaggacca	gctcacagga	aacttcatct	5940
	tcttcgaagt	gtggatttga	tgctcctgg	gtagaaatgt	aggatcttca	aaagtgggcc	6000
	agcctcctgc	acttctctca	aagtctcgcc	tccccaaagg	gtcttaatag	tgctggatgc	6060
	tagctgagtt	agcatcttca	gatgaagagt	aaccctaaag	ttactcttca	gttgccctaa	6120
	ggtagggatgg	tcaactggaa	agctttaaat	taagtccagc	ctaccttggg	ggaaccacc	6180
25	cccacaaaga	aagctgaggt	ccctcctgat	gacttgtcag	tttaactacc	aataaccac	6240
	ttgaattaat	catcatcatc	aagtctttga	taggtgtgag	tgggtatcag	tggccggtcc	6300
	cttcctgggg	ctccagcccc	cgaggaggcc	tcagttagcc	cctgcagaaa	atccatgcat	6360
	catgagtgtc	tcaggggccc	gaatatgaga	gcaggtagga	aacagagaca	tcttccatcc	6420
	ctgagaggca	gtgcggtcca	gtgggtgggg	acacgggctc	tgggtcaggt	ttgtgttgtt	6480
30	tgtttggttg	ttttgagaca	gagtctcgct	ctattgcccc	ggctggagtg	cagtgtcaca	6540
	atctcggtt	actgcaactt	ctgccttccc	ggattcaagt	gattctcctg	cctcagcctc	6600
	cagagtagct	gggattacag	gtgctgcca	ccacgcctgg	ctaatttttg	tatttttgat	6660
	agagacgggg	tttcaccatg	ttggccaggc	tagtctcgaa	ctcttgacct	caagtgatct	6720
	gcctgcctcg	gcctcccaaa	gtgctgggat	tacaggcggtg	agccaccaca	cccagcccca	6780
35	ggttgggtgtt	tgaatctgag	gagactgaag	caccaagggg	ttaaagtgtt	tgcccacagc	6840
	catacttggg	ctcagttcct	tgccctaccc	ctcacttgag	ctgcttagaa	cctgggtgggc	6900



	acatgggcaa	taaccagggtc	acactgtttt	gtaccaagtg	ttatgggaat	ccaagatagg	6960
	agtaatttgc	tctgtggagg	ggatgagggg	tagtgggttag	ggaaagcttc	acaaagtggg	7020
	tgttgcttag	agattttcca	ggtggagaag	ggggcttcta	ggcagaaggc	atagcccaag	7080
	caaagactgc	aagtgcattg	ctgctcatgg	gtagaagaga	atccaccatt	cctcaacatg	7140
5	taccgagtc	ttgccatgtg	caaggcaaca	tgggggtacc	aggaattcca	agcaatgtcc	7200
	aaacctaggg	tctgctttct	gggacctgaa	gatacaggat	ggatcagccc	aggctgcaat	7260
	cccattacca	cgagggggaa	aaaaacctga	aggctaaatt	gtaggctggg	ttagagggtta	7320
	tttatggaaa	gttatattct	acctacatgg	ggtctataag	cctggcgcca	atcagaaaag	7380
	gaacaaacaa	cagacctagc	tgggaggggc	agcattttgt	tgtagggggc	ggggcacatg	7440
10	ttctgggggt	acagccagac	tcagggtctg	tattaatagt	ctgagagtaa	gacagacaga	7500
	gggatagaag	gaaataggct	cctttctctc	tctctctctc	tctctctctc	actctctctc	7560
	tctctcacac	acacacacag	acacacacac	acgtctgtga	ggggctctact	tatgctccaa	7620
	gtacaaatca	ggccacattt	acacaaggag	gtaaaggaaa	agaacgttgg	aggagccaca	7680
	ggaccccaaa	attccctgtt	ttccttgaat	caggcaggac	ttacgcagct	gggaggggtg	7740
15	agagcctgca	gaagccacct	gcgagtaagc	caagttcaga	gtcacagaca	ccaaaagctg	7800
	gtgccatgtc	ccacacccgc	ccacctccca	cctgctcctt	gacacagccc	tgtgctccac	7860
	aaccgggctc	ccagatcatt	gattatagct	ctggggcctg	caccgtcctt	cctgccacat	7920
	ccccacccca	ttcttggaac	ctgccctctg	tcttctccct	tgtccaaggg	caggcaaggg	7980
	ctcagctatt	gggcagcttt	gaccaacagc	tgaggctcct	tttgtggctg	gagatgcagg	8040
20	aggcagggga	atattcctct	tagtcaatgc	gaccatgtgc	ctggtttgcc	cagggtgggc	8100
	tcgtttacac	ctgtaggcca	agcgtaatta	ttaacagctc	ccacttctac	tctaaaaaat	8160
	gacccaatct	gggcagtaaa	ttatatgggtg	cccatgctat	taagagctgc	aacttgctgg	8220
	gcgtgggtgg	tcacacctgt	aatcccagta	ctttgggacg	tcaaggcggg	tggatcacct	8280
	gaggtcacga	gttagagact	ggcctggcca	gcatggcaaa	accccatctt	tactaaaaat	8340
25	acaaaaatta	gcaaggcatg	gtggcatgca	cctgtaatcc	cagggtactcg	ggaggctgag	8400
	acaggagaat	ggcttgaacc	caggaggcag	aggttgacgt	gagccaagat	tgtgccactg	8460
	ccctccagcc	ctggcaacag	agcaagactt	catctcaaaa	gaaaaaggat	actgtcaatc	8520
	actgcaggaa	gaaccacagg	aatgaatgag	gagaagagag	gggctgagtc	accatagtgg	8580
	cagcaccgac	tcctgcagga	aaggcgagac	actgggtcat	gggtactgaa	gggtgccctg	8640
30	aatgacgttc	tgcttttagag	accgaacctg	agccctgaaa	gtgcatgcct	gttcatgggt	8700
	gagagactaa	attcatcatt	ccttggcagg	tactgaatcc	tttcttacgg	ctgccctcca	8760
	atgcccaatt	tccttacaat	tgtctggggg	gcctaagctt	ctgccaccca	agagggccag	8820
	agctggcagc	gagcagctgc	aggtaggaga	gataggatcc	cataagggag	gtgggaaaga	8880
	gagatggaag	gagaggggtg	cagagcacac	acctccctg	cctgacaact	tcctgagggc	8940
35	tgggtcatgcc	agcagattta	aggcggaggc	aggggagatg	gggcgggaga	ggaagtgaag	9000
	aaggagaggg	tggggatgga	gaggaagaga	gggtgatcat	tcattcattc	cattgctact	9060

gactggatgc cagctgtgag ccaggcacca ccctagctct gggcatgtgg ttgtaatctt 9120  
 ggagcctcat ggagctcaca gggagtgtg gcaaggagat ggataatgga cggataacaa 9180  
 ataaacattt agtacaatgt ccgggaatgg aaagttctcg aaagaaaaat aaagctggtg 9240  
 agcatataga cagccctgaa ggcggccagg ccaggcattt ctgaggaggt ggcatttgag 9300  
 5 c 9301

<210> 19

<211> 21

<212> DNA

10 <213> Artificial Sequence

<220>

<223> Primer for PCR

15 <400> 19

ccggagctgg agaacaacaa g 21

<210> 20

<211> 19

20 <212> DNA

<213> Artificial Sequence

<220>

<223> PRimer for PCR

25

<400> 20

gcactggccg gagcacacc 19

<210> 21

30 <211> 23

<212> DNA

<213> Artificial Sequence

<220>

35 <223> Primer for PCR

&lt;400&gt; 21

aggccaaccg cgagaagatg acc

23

&lt;210&gt; 22

5

&lt;211&gt; 21

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

10

&lt;223&gt; Primer for PCR

&lt;400&gt; 22

gaagtccagg gcgacgtagc a

21

15

&lt;210&gt; 23

&lt;211&gt; 25

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

20

&lt;220&gt;

&lt;223&gt; Primer for PCR

&lt;400&gt; 23

aagcttggtgta ccatgcagct cccac

25

25

&lt;210&gt; 24

&lt;211&gt; 50

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

30

&lt;220&gt;

&lt;223&gt; Primer for PCR

&lt;400&gt; 24

35

aagcttctac ttgtcatcgt cgctccttgta gtcgtaggcg ttctccagct

50

<210> 25  
<211> 19  
<212> DNA  
<213> Artificial Sequence

5

<220>  
<223> Primer for PCR

&lt;400&gt; 25

10 gcactggccg gagcacacc

19

<210> 26  
<211> 39  
<212> DNA  
<213> Artificial Sequence

15

<220>  
<223> Primer for PCR

&lt;400&gt; 26

20

gtcgtcggat ccatgggggtg gcaggcggtc aagaatgat

39

<210> 27  
<211> 57  
<212> DNA  
<213> Artificial Sequence

25

<220>  
<223> Primer for PCR

30

&lt;400&gt; 27

gtcgtcaagc ttctacttgt catcgctcctt gtagtcgtag gcgttctcca gctcggc

57

<210> 28  
<211> 29  
<212> DNA

35

<213> Artificial Sequence

<220>

<223> Primer for PCR

5

<400> 28

gacttgatc ccaggggtgg caggcggtc

29

<210> 29

10

<211> 29

<212> DNA

<213> Artificial Sequence

<220>

15

<223> Primer for PCR

<400> 29

agcataagct tctagtaggc gttctccag

29

20

<210> 30

<211> 29

<212> DNA

<213> Artificial Sequence

25

<220>

<223> Primer for PCR

<400> 30

gacttgatc cgaagggaag aagaaagg

29

30

<210> 31

<211> 29

<212> DNA

<213> Artificial Sequence

35

<220>

<223> Primer for PCR

<400> 31

agcataagct tttaatccaa atcgatgga

29

5

<210> 32

<211> 33

<212> DNA

<213> Artificial Sequence

10

<220>

<223> Primer for PCR

<400> 32

actacgagct cggccccacc acccatcaac aag

33

15

<210> 33

<211> 34

<212> DNA

<213> Artificial Sequence

20

<220>

<223> Primer for PCR

<400> 33

acttagaagc tttagtcct cagccccctc ttcc

34

25

<210> 34

<211> 66

<212> DNA

<213> Artificial Sequence

30

<220>

<223> Primer for PCR

35

<400> 34

aatctggatc cataacttcg tatagcatac attatacgaa gttatctgca ggattcgagg 60  
gccccct 66

5 <210> 35  
<211> 82  
<212> DNA  
<213> Artificial Sequence

10 <220>  
<223> Primer for PCR

<400> 35  
aatctgaatt ccaccggtgt taattaaata acttcgtata atgtatgcta tacgaagtta 60  
tagatctaga gtcagcttct ga 82

15 <210> 36  
<211> 62  
<212> DNA  
<213> Artificial Sequence

20 <220>  
<223> Primer for PCR

<400> 36  
25 atttaggtga cactatagaa ctcgagcagc tgaagcttaa ccacatggtg gtcacaacc 60  
at 62

30 <210> 37  
<211> 54  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Primer for PCR

35 <400> 37

aacgacggcc agtgaatccg taatcatggt catgctgcca ggtggaggag ggca

54

<210> 38

<211> 31 <212> DNA

5 <213> Artificial Sequence

<220>

<223> Primer for PCR

10 <400> 38

attaccaccg gtgacacccg cttcctgaca g

31

<210> 39

<211> 61

15 <212> DNA

<213> Artificial Sequence

<220>

<223> Primer for PCR

20

<400> 39

attacttaat taaacatggc gcgccatatg gccggcccct aattgcggcg catcgtaat  
t

60

61

25 <210> 40

<211> 34

<212> DNA

<213> Artificial Sequence

30 <220>

<223> Primer for PCR

<400> 40

attacggccg gccgcaaagg aattcaagat ctga

34

35

<210> 41



44

&lt;211&gt; 34

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

5

&lt;220&gt;

&lt;223&gt; Primer for PCR

&lt;400&gt; 41

attacggcgc gccctcaca ggccgcaccc agct

34

# INTERNATIONAL SEARCH REPORT

International Application No  
PCT/US 99/27990

## A. CLASSIFICATION OF SUBJECT MATTER

IPC 7 C12N15/12 C07K14/51 C07K14/495 C12N15/63 C12N5/10  
C07K16/22 C12Q1/68 C12N15/62 A61K38/18 A61P19/10  
G01N33/53 A01K67/027

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)  
IPC 7 C07K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	BIRREN ET AL.: "Homo sapiens chromosome 17, clone HRPC905N1, complete sequence" EMBL SEQUENCE DATABASE, 14 November 1997 (1997-11-14), XP002133385 HEIDELBERG DE Ac AC003098 the whole document	1,2, 27-30
X	HILLIER ET AL.: "WshU-Merck EST Project 1997" EMBL SEQUENCE DATABASE, 19 May 1997 (1997-05-19), XP002133386 HEIDELBERG DE Ac AA393939 the whole document	1,2, 27-30
	-/--	

☒ Further documents are listed in the continuation of box C.

☒ Patent family members are listed in annex.

### \* Special categories of cited documents:

- "A" document defining the general state of the art which is not considered to be of particular relevance
- "E" earlier document but published on or after the international filing date
- "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- "O" document referring to an oral disclosure, use, exhibition or other means
- "P" document published prior to the international filing date but later than the priority date claimed

- "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
- "&" document member of the same patent family

Date of the actual completion of the international search

20 March 2000

Date of mailing of the international search report

07/04/2000

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2  
NL - 2280 HV Rijswijk  
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,  
Fax: (+31-70) 340-3016

Authorized officer

Ceder, O

# INTERNATIONAL SEARCH REPORT

Int. Application No

PCT/US 99/27990

## C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	<p>BONALDO ET AL.: "Normalization and subtraction: two approaches to facilitate gene discovery" EMBL SEQUENCE DATABASE, 4 September 1998 (1998-09-04), XP002133484 HEIDELBERG DE Ac A1113131 the whole document &amp; BONALDO ET AL.: "Normalization and subtraction: two approaches to facilitate gene discovery" GENOME RES, vol. 6, no. 9, 1996, pages 791-806,</p>	1,27-30
A	<p>US 5 780 263 A (ADAMS MARK D ET AL) 14 July 1998 (1998-07-14) cited in the application column 1, line 11 - line 13 column 1, line 40 - line 42 column 1, line 66 -column 2, line 47 column 9, line 50 - line 53 column 11, line 15 - line 37</p>	1-22,32, 61-67, 73-79
A	<p>US 5 453 492 A (BUETZOW RALF ET AL) 26 September 1995 (1995-09-26)  abstract column 3, line 60 -column 8, line 30</p>	1-3,8,9, 11-13, 15-22, 59,61-67
A	<p>WO 91 13152 A (LUDWIG INST CANCER RES) 5 September 1991 (1991-09-05)  the whole document</p>	1-3,8, 11,13, 15,17,32
A	<p>HSU D R ET AL: "The Xenopus dorsalizing factor Gremlin identifies a novel family of secretes proteins that antagonize BMP activities" MOLECULAR CELL,US,CELL PRESS, CAMBRIDGE, MA, vol. 1, no. 5, April 1998 (1998-04), pages 673-683, XP002113640 ISSN: 1097-2765 cited in the application abstract page 676, left-hand column, line 10 - line 14</p>	17
A	<p>WO 92 06693 A (FOX CHASE CANCER CENTER) 30 April 1992 (1992-04-30)</p>	

## INTERNATIONAL SEARCH REPORT

International application No.

PCT/US 99/ 27990

### Box I Observations where certain claims were found unsearchable (Continuation of Item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☒ Claims Nos.: 57 and 58  
because they relate to subject matter not required to be searched by this Authority, namely:  
see PCT/ISA/210
2. ☐ Claims Nos.:  
because they relate to parts of the International Application that do not comply with the prescribed requirements to such  
an extent that no meaningful International Search can be carried out, specifically:
3. ☐ Claims Nos.:  
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

### Box II Observations where unity of invention is lacking (Continuation of Item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. ☐ As all required additional search fees were timely paid by the applicant, this International Search Report covers all  
searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment  
of any additional fee.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this International Search Report  
covers only those claims for which fees were paid, specifically claims Nos.:
4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is  
restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

#### Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest.
- ☐ No protest accompanied the payment of additional search fees.

# INTERNATIONAL SEARCH REPORT

International Application No. PCT/US 99 27990

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

Continuation of Box I.1

Although claims 57 and 58 are directed to a method of treatment of the human/animal body, the search has been carried out and based on the alleged effects of the compound/composition.

---

Continuation of Box I.1

Rule 39.1(iv) PCT - Method for treatment of the human or animal body by therapy

# INTERNATIONAL SEARCH REPORT

information on patent family members

International Application No

PCT/US 99/27990

Patent document cited in search report		Publication date	Patent family member(s)	Publication date
US 5780263	A	14-07-1998	CA 2220912 A	12-12-1996
			WO 9639486 A	12-12-1996
			AU 2766595 A	24-12-1996
			EP 0871705 A	21-10-1996
			JP 11506918 T	22-06-1999
US 5453492	A	26-09-1995	NONE	
WO 9113152	A	05-09-1991	US 5177197 A	05-01-1993
			AU 649026 B	12-05-1994
			AU 7449591 A	18-09-1991
			CA 2076979 A	28-08-1991
			DE 69131572 D	07-10-1999
			DE 69131572 T	23-12-1999
			EP 0517779 A	16-12-1992
			JP 5504888 T	29-07-1993
WO 9206693	A	30-04-1992	AU 662304 B	31-08-1995
			AU 8957591 A	20-05-1992
			CA 2094608 A	23-04-1992
			EP 0554376 A	11-08-1993
			JP 6502311 T	17-03-1994

## CORRECTED VERSION

(19) World Intellectual Property Organization  
International Bureau(43) International Publication Date  
8 June 2000 (08.06.2000)

PCT

(10) International Publication Number  
WO 00/32773 A1(51) International Patent Classification<sup>7</sup>: C12N 15/12,  
C07K 14/51, 14/495, C12N 15/63, 5/10, C07K 16/22,  
C12Q 1/68, C12N 15/62, A61K 38/18, A61P 19/10, G01N  
33/53, A01K 67/027(71) Applicant (for all designated States except US): DARWIN  
DISCOVERY LTD. [GB/GB]; Cambridge Science Park,  
Milton Road, Cambridge, Cambridgeshire CB4 4WE (GB).

(21) International Application Number: PCT/US99/27990

(72) Inventors; and

(22) International Filing Date:

24 November 1999 (24.11.1999)

(75) Inventors/Applicants (for US only): BRUNKOW, Mary,  
E. [US/US]; 9829 Triton Drive NW, Seattle, WA 98117  
(US). GALAS, David, J. [US/US]; 854 Guanajuato Drive,  
Claremont, CA 91711 (US). KOVACEVICH, Brian  
[US/US]; 4308 N.E. 6th Place, Renton, WA 98059 (US).  
MULLIGAN, John, T. [US/US]; 5823 17th Avenue  
Northeast, Seattle, WA 98105 (US). PAEPER, Bryan, W.  
[US/US]; 1617 Summit Avenue #43, Seattle, WA 98122  
(US). VAN NESS, Jeffrey [US/US]; 10020 49th Avenue  
Northeast, Seattle, WA 98125 (US). WINKLER, David,  
G. [US/US]; 7037 20th Avenue NE, Seattle, WA 98115  
(US).

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

60/110,283 27 November 1998 (27.11.1998) US

[Continued on next page]

(54) Title: COMPOSITIONS AND METHODS FOR INCREASING BONE MINERALIZATION

## Common Cysteine Backbone

1	50
human-gremlin pro	_____
human-cerberus pro	PHLLLFGLLV LLPLGKTTTH QDGRNDSSL SPYLLPRHOR ELPTGAKCEA
human-dan pro	_____
human-beer pro	_____
51	100
human-gremlin pro	_____H SRTAYTVGAL LLLLGTLPLA AEGKKKGSQD
human-cerberus pro	EKPDLFVAV PHLVAT SPA GEGGQDREKM LSRGFRNWK PEREMHPSRD
human-dan pro	_____
human-beer pro	_____MDLPLA LGLVLLVHT
101	150
human-gremlin pro	AI PPKQKAD HNSGQTSQSP QDPGSRNRP GEGGRTAMPS EGVLESSCEA
human-cerberus pro	SDSEPFPPGT QSLQIPID G MUNEKSPLE EAKKFAHFM FRKTPASQGY
human-dan pro	_____MRLVVGAVL PAMLLAAPP
human-beer pro	AFRYVEGGGA QAFKQDQTEI TPELGEVPEP PPELENMTH KRAKGGRRP
151	200
human-gremlin pro	LHVTERKYLK RQWCKTQPLK QTIHEGGNS RTIINRT CY GQCKSYIFFR
human-cerberus pro	ILPIKSHVH WETGRVPS QTIHEGCEK VVYQNAL CF GQCKSYHFF
human-dan pro	IMKALFPDK SAWCAKMTT QVWHSCEA KSIDNRA CL GQCKSYSPVH
human-beer pro	HMPETKQVS EYSCEHLHT RYVTDGERS AKPVTELVES GQCKPARLLP
201	250
human-gremlin pro	HTRKEEGSPG SCSE CKP KKEFTTHMYL AEPELOPPTK K KRYTRVQ
human-cerberus pro	GAADHSHT SCSE CLP AKFTTHMLPL NTELSVVK V VNLVEE
human-dan pro	TFPOSTESLY HDS EMP AOSHWELVTL ECPGHEEVR VDKLVKILH
human-beer pro	NAIGRQWAK PSQDFRCLP DEYRADRVOL LCPGGEAPRA RNVLYVAS
251	300
human-gremlin pro	GRG ISIDLG
human-cerberus pro	COCKVTEHE GGHLEHSSQ DSFIPGWSA
human-dan pro	CSCQACQKEP SHELSYVQ GEDGSSQFG THPHPHPHH PGQGTPEPLD
human-beer pro	CKCKRL TRFH HQSELKDFGT EAARPQDGA PPRAARSACA HQALEENAY
301	314
human-gremlin pro	_____
human-cerberus pro	_____
human-dan pro	PPGAPHTTEE GAED
human-beer pro	_____

(57) Abstract: A novel class or family of TGF- $\beta$  binding proteins is disclosed. Also disclosed are assays for selecting molecules for increasing bone mineralization and methods for utilizing such molecules.

WO 00/32773 A1



(74) Agent: MCMASTERS, David, D.: Seed and Berry LLP,  
Suite 6300, 701 Fifth Avenue, Seattle, WA 98104-7092  
(US).

MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM,  
GA, GN, GW, ML, MR, NE, SN, TD, TG).

**Published:**

— with international search report

(81) Designated States (*national*): AE, AL, AM, AT, AU, AZ,  
BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK,  
DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL,  
IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU,  
LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT,  
RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA,  
UG, US, UZ, VN, YU, ZA, ZW.

(48) Date of publication of this corrected version:

13 December 2001

(15) Information about Correction:

see PCT Gazette No. 50/2001 of 13 December 2001, Sec-  
tion II

(84) Designated States (*regional*): ARIPO patent (GH, GM,  
KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent  
(AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent  
(AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU,

*For two-letter codes and other abbreviations, refer to the "Guid-  
ance Notes on Codes and Abbreviations" appearing at the begin-  
ning of each regular issue of the PCT Gazette.*



## COMPOSITIONS AND METHODS FOR INCREASING BONE MINERALIZATION

### TECHNICAL FIELD

The present invention relates generally to pharmaceutical products and methods and, more specifically, to methods and compositions suitable for increasing the mineral content of bone. Such compositions and methods may be utilized to treat a wide variety of conditions, including for example, osteopenia, osteoporosis, fractures and other disorders in which low bone mineral density are a hallmark of the disease.

### BACKGROUND OF THE INVENTION

Two or three distinct phases of changes to bone mass occur over the life of an individual (see Riggs, *West J. Med.* 154:63-77, 1991). The first phase occurs in both men and women, and proceeds to attainment of a peak bone mass. This first phase is achieved through linear growth of the endochondral growth plates, and radial growth due to a rate of periosteal apposition. The second phase begins around age 30 for trabecular bone (flat bones such as the vertebrae and pelvis) and about age 40 for cortical bone (e.g., long bones found in the limbs) and continues to old age. This phase is characterized by slow bone loss, and occurs in both men and women. In women, a third phase of bone loss also occurs, most likely due to postmenopausal estrogen deficiencies. During this phase alone, women may lose an additional 10% of bone mass from the cortical bone and 25% from the trabecular compartment (see Riggs, *supra*).

Loss of bone mineral content can be caused by a wide variety of conditions, and may result in significant medical problems. For example, osteoporosis is a debilitating disease in humans characterized by marked decreases in skeletal bone mass and mineral density, structural deterioration of bone including degradation of bone microarchitecture and corresponding increases in bone fragility and susceptibility to fracture in afflicted individuals. Osteoporosis in humans is preceded by clinical osteopenia (bone mineral density that is greater than one standard deviation but less than 2.5 standard deviations below the mean value for young adult bone), a condition found in approximately 25 million people in the United States. Another 7-8 million patients in the United States have been diagnosed with clinical osteoporosis (defined as bone mineral content greater than 2.5 standard deviations below that of mature young adult bone). Osteoporosis is one of the most expensive diseases for the health care

system, costing tens of billions of dollars annually in the United States. In addition to health care-related costs, long-term residential care and lost working days add to the financial and social costs of this disease. Worldwide approximately 75 million people are at risk for osteoporosis.

5           The frequency of osteoporosis in the human population increases with age, and among Caucasians is predominant in women (who comprise 80% of the osteoporosis patient pool in the United States). The increased fragility and susceptibility to fracture of skeletal bone in the aged is aggravated by the greater risk of  
10 fractures are reported in the United States each year. Fractured hips, wrists, and vertebrae are among the most common injuries associated with osteoporosis. Hip fractures in particular are extremely uncomfortable and expensive for the patient, and for women correlate with high rates of mortality and morbidity.

15           Although osteoporosis has been defined as an increase in the risk of fracture due to decreased bone mass, none of the presently available treatments for skeletal disorders can substantially increase the bone density of adults. There is a strong perception among all physicians that drugs are needed which could increase bone density in adults, particularly in the bones of the wrist, spinal column and hip that are at risk in osteopenia and osteoporosis.

20           Current strategies for the prevention of osteoporosis may offer some benefit to individuals but cannot ensure resolution of the disease. These strategies include moderating physical activity (particularly in weight-bearing activities) with the onset of advanced age, including adequate calcium in the diet, and avoiding consumption of products containing alcohol or tobacco. For patients presenting with  
25 clinical osteopenia or osteoporosis, all current therapeutic drugs and strategies are directed to reducing further loss of bone mass by inhibiting the process of bone absorption, a natural component of the bone remodeling process that occurs constitutively.

30           For example, estrogen is now being prescribed to retard bone loss. There is, however, some controversy over whether there is any long term benefit to patients and whether there is any effect at all on patients over 75 years old. Moreover, use of estrogen is believed to increase the risk of breast and endometrial cancer.

35           High doses of dietary calcium, with or without vitamin D has also been suggested for postmenopausal women. However, high doses of calcium can often have unpleasant gastrointestinal side effects, and serum and urinary calcium levels must be continuously monitored (see Khosla and Riggs, *Mayo Clin. Proc.* 70:978-982, 1995).

Other therapeutics which have been suggested include calcitonin, bisphosphonates, anabolic steroids and sodium fluoride. Such therapeutics however, have undesirable side effects (e.g., calcitonin and steroids may cause nausea and provoke an immune reaction, bisphosphonates and sodium fluoride may inhibit repair of fractures, even though bone density increases modestly) that may prevent their usage (see Khosla and Riggs, *supra*).

No currently practiced therapeutic strategy involves a drug that stimulates or enhances the growth of new bone mass. The present invention provides compositions and methods which can be utilized to increase bone mineralization, and thus may be utilized to treat a wide variety of conditions where it is desired to increase bone mass. Further, the present invention provides other, related advantages.

## SUMMARY OF THE INVENTION

As noted above, the present invention provides a novel class or family of TGF-beta binding-proteins, as well as assays for selecting compounds which increase bone mineral content and bone mineral density, compounds which increase bone mineral content and bone mineral density and methods for utilizing such compounds in the treatment or prevention of a wide variety of conditions.

Within one aspect of the present invention, isolated nucleic acid molecules are provided, wherein said nucleic acid molecules are selected from the group consisting of: (a) an isolated nucleic acid molecule comprising sequence ID Nos. 1, 5, 7, 9, 11, 13, or, 15, or complementary sequence thereof; (b) an isolated nucleic acid molecule that specifically hybridizes to the nucleic acid molecule of (a) under conditions of high stringency; and (c) an isolated nucleic acid that encodes a TGF-beta binding-protein according to (a) or (b). Within related aspects of the present invention, isolated nucleic acid molecules are provided based upon hybridization to only a portion of one of the above-identified sequences (e.g., for (a) hybridization may be to a probe of at least 20, 25, 50, or 100 nucleotides selected from nucleotides 156 to 539 or 555 to 687 of Sequence ID No. 1). As should be readily evident, the necessary stringency to be utilized for hybridization may vary based upon the size of the probe. For example, for a 25-mer probe high stringency conditions could include: 60 mM Tris pH 8.0, 2 mM EDTA, 5x Denhardt's, 6x SSC, 0.1% (w/v) N-laurylsarcosine, 0.5% (w/v) NP-40 (nonidet P-40) overnight at 45 degrees C, followed by two washes with 0.2x SSC / 0.1% SDS at 45-50 degrees. For a 100-mer probe under low stringency conditions, suitable conditions might include the following: 5x SSPE, 5x Denhardt's, and 0.5% SDS overnight at 42-50 degrees, followed by two washes with 2x SSPE (or 2x SSC)

10.1% SDS at 42-50 degrees.

Within related aspects of the present invention, isolated nucleic acid molecules are provided which have homology to Sequence ID Nos. 1, 5, 7, 9, 11, 13, or 15, at a 50%, 60%, 75%, 80%, 90%, 95%, or 98% level of homology utilizing a Wilbur-Lipman algorithm. Representative examples of such isolated molecules include, for example, nucleic acid molecules which encode a protein comprising Sequence ID NOs. 2, 6, 10, 12, 14, or 16, or have homology to these sequences at a level of 50%, 60%, 75%, 80%, 90%, 95%, or 98% level of homology utilizing a Lipman-Pearson algorithm.

Isolated nucleic acid molecules are typically less than 100kb in size, and, within certain embodiments, less than 50kb, 25kb, 10kb, or even 5kb in size. Further, isolated nucleic acid molecules, within other embodiments, do not exist in a "library" of other unrelated nucleic acid molecules (*e.g.*, a subclone BAC such as described in GenBank Accession No. AC003098 and EMB No. AQ171546). However, isolated nucleic acid molecules can be found in libraries of related molecules (*e.g.*, for shuffling, such as is described in U.S. Patent Nos. 5,837,458; 5,830,721; and 5,811,238). Finally, isolated nucleic acid molecules as described herein do not include nucleic acid molecules which encode Dan, Cerberus, Gremlin, or SCGF (U.S. Patent No. 5,780,263).

Also provided by the present invention are cloning vectors which contain the above-noted nucleic acid molecules, and expression vectors which comprise a promoter (*e.g.*, a regulatory sequence) operably linked to one of the above-noted nucleic acid molecules. Representative examples of suitable promoters include tissue-specific promoters, and viral - based promoters (*e.g.*, CMV-based promoters such as CMV I-E, SV40 early promoter, and MuLV LTR). Expression vectors may also be based upon, or derived from viruses (*e.g.*, a "viral vector"). Representative examples of viral vectors include herpes simplex viral vectors, adenoviral vectors, adenovirus-associated viral vectors and retroviral vectors. Also provided are host cells containing or comprising any of above-noted vectors (including for example, host cells of human, monkey, dog, rat, or mouse origin).

Within other aspects of the present invention, methods of producing TGF-beta binding-proteins are provided, comprising the step of culturing the aforementioned host cell containing vector under conditions and for a time sufficient to produce the TGF-beta binding protein. Within further embodiments, the protein produced by this method may be further purified (*e.g.*, by column chromatography, affinity purification, and the like). Hence, isolated proteins which are encoded by the

above-noted nucleic acid molecules (*e.g.*, Sequence ID NOs. 2, 4, 6, 8, 10, 12, 14, or 16) may be readily produced given the disclosure of the subject application.

It should also be noted that the aforementioned proteins, or fragments thereof, may be produced as fusion proteins. For example, within one aspect fusion proteins are provided comprising a first polypeptide segment comprising a TGF-beta binding-protein encoded by a nucleic acid molecule as described above, or a portion thereof of at least 10, 20, 30, 50, or 100 amino acids in length, and a second polypeptide segment comprising a non-TGF-beta binding-protein. Within certain embodiments, the second polypeptide may be a tag suitable for purification or recognition (*e.g.*, a polypeptide comprising multiple anionic amino acid residues – see U.S. Patent No. 4,851,341), a marker (*e.g.*, green fluorescent protein, or alkaline phosphatase), or a toxic molecule (*e.g.*, ricin).

Within another aspect of the present invention, antibodies are provided which are capable of specifically binding the above-described class of TGF-beta binding proteins (*e.g.*, human BEER). Within various embodiments, the antibody may be a polyclonal antibody, or a monoclonal antibody (*e.g.*, of human or murine origin). Within further embodiments, the antibody is a fragment of an antibody which retains the binding characteristics of a whole antibody (*e.g.*, an F(ab')<sub>2</sub>, F(ab)<sub>2</sub>, Fab', Fab, or Fv fragment, or even a CDR). Also provided are hybridomas and other cells which are capable of producing or expressing the aforementioned antibodies.

Within related aspects of the invention, methods are provided detecting a TGF-beta binding protein, comprising the steps of incubating an antibody as described above under conditions and for a time sufficient to permit said antibody to bind to a TGF-beta binding protein, and detecting the binding. Within various embodiments the antibody may be bound to a solid support to facilitate washing or separation, and/or labeled. (*e.g.*, with a marker selected from the group consisting of enzymes, fluorescent proteins, and radioisotopes).

Within other aspects of the present invention, isolated oligonucleotides are provided which hybridize to a nucleic acid molecule according to Sequence ID NOs. 1, 3, 5, 7, 9, 11, 13, 15, 17, or 18 or the complement thereto, under conditions of high stringency. Within further embodiments, the oligonucleotide may be found in the sequence which encodes Sequence ID Nos. 2, 4, 6, 8, 10, 12, 14, or 16. Within certain embodiments, the oligonucleotide is at least 15, 20, 30, 50, or 100 nucleotides in length. Within further embodiments, the oligonucleotide is labeled with another molecule (*e.g.*, an enzyme, fluorescent molecule, or radioisotope). Also provided are primers which are capable of specifically amplifying all or a portion of the above-

mentioned nucleic acid molecules which encode TGF-beta binding-proteins. As utilized herein, the term "specifically amplifying" should be understood to refer to primers which amplify the aforementioned TGF-beta binding-proteins, and not other TGF-beta binding proteins such as Dan, Cerberus, Gremlin, or SCGF (U.S. Patent No. 5,780,263).

Within related aspects of the present invention, methods are provided for detecting a nucleic acid molecule which encodes a TGF-beta binding protein, comprising the steps of incubating an oligonucleotide as described above under conditions of high stringency, and detecting hybridization of said oligonucleotide. Within certain embodiments, the oligonucleotide may be labeled and/or bound to a solid support.

Within other aspects of the present invention, ribozymes are provided which are capable of cleaving RNA which encodes one of the above-mentioned TGF-beta binding-proteins (e.g., Sequence ID NOs. 2, 6, 8, 10, 12, 14, or 16). Such ribozymes may be composed of DNA, RNA (including 2'-O-methyl ribonucleic acids), nucleic acid analogs (e.g., nucleic acids having phosphorothioate linkages) or mixtures thereof. Also provided are nucleic acid molecules (e.g., DNA or cDNA) which encode these ribozymes, and vectors which are capable of expressing or producing the ribozymes. Representative examples of vectors include plasmids, retrotransposons, cosmids, and viral-based vectors (e.g., viral vectors generated at least in part from a retrovirus, adenovirus, or, adeno-associated virus). Also provided are host cells (e.g., human, dog, rat, or mouse cells) which contain these vectors. In certain embodiments, the host cell may be stably transformed with the vector.

Within further aspects of the invention, methods are provided for producing ribozymes either synthetically, or by *in vitro* or *in vivo* transcription. Within further embodiments, the ribozymes so produced may be further purified and / or formulated into pharmaceutical compositions (e.g., the ribozyme or nucleic acid molecule encoding the ribozyme along with a pharmaceutically acceptable carrier or diluent). Similarly, the antisense oligonucleotides and antibodies or other selected molecules described herein may be formulated into pharmaceutical compositions.

Within other aspects of the present invention, antisense oligonucleotides are provided comprising a nucleic acid molecule which hybridizes to a nucleic acid molecule according to Sequence ID NOs. 1, 3, 5, 7, 9, 11, 13, or 15, or the complement thereto, and wherein said oligonucleotide inhibits the expression of TGF-beta binding protein as described herein (e.g., human BEER). Within various embodiments, the oligonucleotide is 15, 20, 25, 30, 35, 40, or 50 nucleotides in length. Preferably, the

oligonucleotide is less than 100, 75, or 60 nucleotides in length. As should be readily evident, the oligonucleotide may be comprised of one or more nucleic acid analogs, ribonucleic acids, or deoxyribonucleic acids. Further, the oligonucleotide may be modified by one or more linkages, including for example, covalent linkage such as a phosphorothioate linkage, a phosphotriester linkage, a methyl phosphonate linkage, a methylene(methylimino) linkage, a morpholino linkage, an amide linkage, a polyamide linkage, a short chain alkyl intersugar linkage, a cycloalkyl intersugar linkage, a short chain heteroatomic intersugar linkage and a heterocyclic intersugar linkage. One representative example of a chimeric oligonucleotide is provided in U.S. Patent No. 5,989,912.

Within yet another aspect of the present invention, methods are provided for increasing bone mineralization, comprising introducing into a warm-blooded animal an effective amount of the ribozyme as described above. Within related aspects, such methods comprise the step of introducing into a patient an effective amount of the nucleic acid molecule or vector as described herein which is capable of producing the desired ribozyme, under conditions favoring transcription of the nucleic acid molecule to produce the ribozyme.

Within other aspects of the invention transgenic, non-human animals are provided. Within one embodiment a transgenic animal is provided whose germ cells and somatic cells contain a nucleic acid molecule encoding a TGF-beta binding-protein as described above which is operably linked to a promoter effective for the expression of the gene, the gene being introduced into the animal, or an ancestor of the animal, at an embryonic stage, with the proviso that said animal is not a human. Within other embodiments, transgenic knockout animals are provided, comprising an animal whose germ cells and somatic cells comprise a disruption of at least one allele of an endogenous nucleic acid molecule which hybridizes to a nucleic acid molecule which encodes a TGF-binding protein as described herein, wherein the disruption prevents transcription of messenger RNA from said allele as compared to an animal without the disruption, with the proviso that the animal is not a human. Within various embodiments, the disruption is a nucleic acid deletion, substitution, or, insertion. Within other embodiments the transgenic animal is a mouse, rat, sheep, pig, or dog.

Within further aspects of the invention, kits are provided for the detection of TGF-beta binding-protein gene expression, comprising a container that comprises a nucleic acid molecule, wherein the nucleic acid molecule is selected from the group consisting of (a) a nucleic acid molecule comprising the nucleotide sequence of SEQ ID NOs: 1, 3, 5, 7, 9, 11, 13, or 15; (b) a nucleic acid molecule comprising the

complement of the nucleotide sequence of (a); (c) a nucleic acid molecule that is a fragment of (a) or (b) of at least 15, 20 30, 50, 75, or, 100 nucleotides in length. Also provided are kits for the detection of a TGF-beta binding-protein which comprise a container that comprise one of the TGF-beta binding protein antibodies described  
5 herein.

For example, within one aspect of the present invention methods are provided for determining whether a selected molecule is capable of increasing bone mineral content, comprising the steps of (a) mixing one or more candidate molecules with TGF-beta-binding-protein encoded by the nucleic acid molecule according to  
10 claim 1 and a selected member of the TGF-beta family of proteins (*e.g.*, BMP 5 or 6), (b) determining whether the candidate molecule alters the signaling of the TGF-beta family member, or alters the binding of the TGF-beta binding-protein to the TGF-beta family member. Within certain embodiments, the molecule alters the ability of TGF-beta to function as a positive regulator of mesenchymal cell differentiation. Within  
15 this aspect of the present invention, the candidate molecule(s) may alter signaling or binding by, for example, either decreasing (*e.g.*, inhibiting), or increasing (*e.g.*, enhancing) signaling or binding.

Within yet another aspect, methods are provided for determining whether a selected molecule is capable of increasing bone mineral content, comprising  
20 the step of determining whether a selected molecule inhibits the binding of TGF-beta binding-protein to bone, or an analogue thereof. Representative examples of bone or analogues thereof include hydroxyapatite and primary human bone samples obtained via biopsy.

Within certain embodiments of the above-recited methods, the selected  
25 molecule is contained within a mixture of molecules and the methods may further comprise the step of isolating one or more molecules which are functional within the assay. Within yet other embodiments, TGF-beta family of proteins is bound to a solid support and the binding of TGF-beta binding-protein is measured or TGF-beta binding-protein are bound to a solid support and the binding of TGF-beta proteins are measured.  
30

Utilizing methods such as those described above, a wide variety of molecules may be assayed for their ability to increase bone mineral content by inhibiting the binding of the TGF-beta binding-protein to the TGF-beta family of proteins. Representative examples of such molecules include proteins or peptides, organic molecules, and nucleic acid molecules.  
35

Within other related aspects of the invention, methods are provided for increasing bone mineral content in a warm-blooded animal, comprising the step of



administering to a warm-blooded animal a therapeutically effective amount of a molecule identified from the assays recited herein. Within another aspect, methods are provided for increasing bone mineral content in a warm-blooded animal, comprising the step of administering to a warm-blooded animal a therapeutically effective amount  
5 of a molecule which inhibits the binding of the TGF-beta binding-protein to the TGF-beta super-family of proteins, including bone morphogenic proteins (BMPs). Representative examples of suitable molecules include antisense molecules, ribozymes, ribozyme genes, and antibodies (e.g., a humanized antibody) which specifically recognize and alter the activity of the TGF-beta binding-protein.

10 Within another aspect of the present invention, methods are provided for increasing bone mineral content in a warm-blooded animal, comprising the steps of (a) introducing into cells which home to the bone a vector which directs the expression of a molecule which inhibits the binding of the TGF-beta binding-protein to the TGF-beta family of proteins and bone morphogenic proteins (BMPs), and (b) administering  
15 the vector-containing cells to a warm-blooded animal. As utilized herein, it should be understood that cells "home to bone" if they localize within the bone matrix after peripheral administration. Within one embodiment, such methods further comprise, prior to the step of introducing, isolating cells from the marrow of bone which home to the bone. Within a further embodiment, the cells which home to bone are selected from  
20 the group consisting of CD34+ cells and osteoblasts.

Within other aspects of the present invention, molecules are provided (preferably isolated) which inhibit the binding of the TGF-beta binding-protein to the TGF-beta super-family of proteins.

Within further embodiments, the molecules may be provided as a  
25 composition, and can further comprise an inhibitor of bone resorption. Representative examples of such inhibitors include calcitonin, estrogen, a bisphosphonate, a growth factor having anti-resorptive activity and tamoxifen.

Representative examples of molecules which may be utilized in the afore-mentioned therapeutic contexts include, e.g., ribozymes, ribozyme genes, antisense molecules, and/or antibodies (e.g., humanized antibodies). Such molecules  
30 may depending upon their selection, used to alter, antagonize, or agonize the signalling or binding of a TGF-beta binding-protein family member as described herein

Within various embodiments of the invention, the above-described molecules and methods of treatment or prevention may be utilized on conditions such  
35 as osteoporosis, osteomalasia, periodontal disease, scurvy, Cushing's Disease, bone fracture and conditions due to limb immobilization and steroid usage.

These and other aspects of the present invention will become evident upon reference to the following detailed description and attached drawings. In addition, various references are set forth herein which describe in more detail certain procedures or compositions (e.g., plasmids, etc.), and are therefore incorporated by  
5 reference in their entirety.

## BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 is a schematic illustration comparing the amino acid sequence of Human Dan; Human Gremlin; Human Cerberus and Human Beer. Arrows indicate the Cysteine backbone.

10 Figure 2 summarizes the results obtained from surveying a variety of human tissues for the expression of a TGF-beta binding-protein gene, specifically, the Human Beer gene. A semi-quantitative Reverse Transcription-Polymerase Chain Reaction (RT-PCR) procedure was used to amplify a portion of the gene from first-strand cDNA synthesized from total RNA (described in more detail in EXAMPLE 2A).

15 Figure 3 summarizes the results obtained from RNA *in situ* hybridization of mouse embryo sections, using a cRNA probe that is complementary to the mouse Beer transcript (described in more detail in EXAMPLE 2B). Panel A is a transverse section of 10.5 dpc embryo. Panel B is a sagittal section of 12.5 dpc embryo and panels C and D are sagittal sections of 15.5 dpc embryos.

20 Figure 4 illustrates, by western blot analysis, the specificity of three different polyclonal antibodies for their respective antigens (described in more detail in EXAMPLE 4). Figure 4A shows specific reactivity of an anti-H. Beer antibody for H. Beer antigen, but not H. Dan or H. Gremlin. Figure 4B shows reactivity of an anti-H. Gremlin antibody for H. Gremlin antigen, but not H. Beer or H. Dan. Figure 4C shows  
25 reactivity of an anti-H. Dan antibody for H. Dan, but not H. Beer or H. Gremlin.

Figure 5 illustrates, by western blot analysis, the selectivity of the TGF-beta binding-protein, Beer, for BMP-5 and BMP-6, but not BMP-4 (described in more detail in EXAMPLE 5).

30 Figure 6 demonstrates that the ionic interaction between the TGF-beta binding-protein, Beer, and BMP-5 has a dissociation constant in the 15-30 nM range.

## DETAILED DESCRIPTION OF THE INVENTION

## DEFINITIONS

Prior to setting forth the invention in detail, it may be helpful to an understanding thereof to set forth definitions of certain terms and to list and to define the abbreviations that will be used hereinafter.

“Molecule” should be understood to include proteins or peptides (*e.g.*, antibodies, recombinant binding partners, peptides with a desired binding affinity), nucleic acids (*e.g.*, DNA, RNA, chimeric nucleic acid molecules, and nucleic acid analogues such as PNA); and organic or inorganic compounds.

“TGF-beta” should be understood to include any known or novel member of the TGF-beta super-family, which also includes bone morphogenic proteins (BMPs).

“TGF-beta receptor” should be understood to refer to the receptor specific for a particular member of the TGF-beta super-family (including bone morphogenic proteins (BMPs)).

“TGF-beta binding-protein” should be understood to refer to a protein with specific binding affinity for a particular member or subset of members of the TGF-beta super-family (including bone morphogenic proteins (BMPs)). Specific examples of TGF-beta binding-proteins include proteins encoded by Sequence ID Nos. 1, 5, 7, 9, 11, 13, and 15.

Inhibiting the “binding of the TGF-beta binding-protein to the TGF-beta family of proteins and bone morphogenic proteins (BMPs)” should be understood to refer to molecules which allow the activation of TGF-beta or bone morphogenic proteins (BMPs), or allow the binding of TGF-beta family members including bone morphogenic proteins (BMPs) to their respective receptors, by removing or preventing TGF-beta from binding to TGF-binding-protein. Such inhibition may be accomplished, for example, by molecules which inhibit the binding of the TGF-beta binding-protein to specific members of the TGF-beta super-family.

“Vector” refers to an assembly which is capable of directing the expression of desired protein. The vector must include transcriptional promoter elements which are operably linked to the gene(s) of interest. The vector may be composed of either deoxyribonucleic acids (“DNA”), ribonucleic acids (“RNA”), or a combination of the two (*e.g.*, a DNA-RNA chimeric). Optionally, the vector may include a polyadenylation sequence, one or more restriction sites, as well as one or more selectable markers such as neomycin phosphotransferase or hygromycin phosphotransferase. Additionally, depending on the host cell chosen and the vector

employed. other genetic elements such as an origin of replication, additional nucleic acid restriction sites, enhancers, sequences conferring inducibility of transcription, and selectable markers, may also be incorporated into the vectors described herein.

An "isolated nucleic acid molecule" is a nucleic acid molecule that is not integrated in the genomic DNA of an organism. For example, a DNA molecule that encodes a TGF-binding protein that has been separated from the genomic DNA of a eukaryotic cell is an isolated DNA molecule. Another example of an isolated nucleic acid molecule is a chemically-synthesized nucleic acid molecule that is not integrated in the genome of an organism. The isolated nucleic acid molecule may be genomic DNA, cDNA, RNA, or composed at least in part of nucleic acid analogs.

An "isolated polypeptide" is a polypeptide that is essentially free from contaminating cellular components, such as carbohydrate, lipid, or other proteinaceous impurities associated with the polypeptide in nature. Within certain embodiments, a particular protein preparation contains an isolated polypeptide if it appears nominally as a single band on SDS-PAGE gel with Coomassie Blue staining. "Isolated" when referring to organic molecules means that the compounds are greater than 90 percent pure utilizing methods which are well known in the art (*e.g.*, NMR, melting point).

"Sclerosteosis" Sclerosteosis is a term that was applied by Hansen (1967) (Hansen, H. G., Sklerosteose. In: Opitz, H.; Schmid, F., Handbuch der Kinderheilkunde. Berlin: Springer (pub.) 6 1967. Pp. 351-355) to a disorder similar to van Buchem hyperostosis corticalis generalisata but possibly differing in radiologic appearance of the bone changes and in the presence of asymmetric cutaneous syndactyly of the index and middle fingers in many cases. The jaw has an unusually square appearance in this condition.

"Humanized antibodies" are recombinant proteins in which murine complementary determining regions of monoclonal antibodies have been transferred from heavy and light variable chains of the murine immunoglobulin into a human variable domain.

As used herein, an "antibody fragment" is a portion of an antibody such as F(ab')<sub>2</sub>, F(ab)<sub>2</sub>, Fab', Fab, and the like. Regardless of structure, an antibody fragment binds with the same antigen that is recognized by the intact antibody. For example, an anti-TGF-beta binding-protein monoclonal antibody fragment binds with an epitope of TGF-beta binding-protein.

The term "antibody fragment" also includes any synthetic or genetically engineered protein that acts like an antibody by binding to a specific antigen to form a complex. For example, antibody fragments include isolated fragments consisting of the

light chain variable region, "Fv" fragments consisting of the variable regions of the heavy and light chains, recombinant single chain polypeptide molecules in which light and heavy variable regions are connected by a peptide linker ("sFv proteins"), and minimal recognition units consisting of the amino acid residues that mimic the hypervariable region.

A "detectable label" is a molecule or atom which can be conjugated to an antibody moiety to produce a molecule useful for diagnosis. Examples of detectable labels include chelators, photoactive agents, radioisotopes, fluorescent agents, paramagnetic ions, enzymes, and other marker moieties.

As used herein, an "immunoconjugate" is a molecule comprising an anti-TGF-beta binding-protein antibody, or an antibody fragment, and a detectable label. An immunoconjugate has roughly the same, or only slightly reduced, ability to bind TGF-beta binding-protein after conjugation as before conjugation.

Abbreviations: TGF-beta - "Transforming Growth Factor-beta"; TGF-bBP - "Transforming Growth Factor-beta binding-protein" (one representative TGF-bBP is designated "H. Beer"); BMP - "bone morphogenic protein"; PCR - "polymerase chain reaction"; RT-PCR - PCR process in which RNA is first transcribed into DNA at the first step using reverse transcriptase (RT); cDNA - any DNA made by copying an RNA sequence into DNA form.

As noted above, the present invention provides a novel class of TGF-beta binding-proteins, as well as methods and compositions for increasing bone mineral content in warm-blooded animals. Briefly, the present inventions are based upon the unexpected discovery that a mutation in the gene which encodes a novel member of the TGF-beta binding-protein family results in a rare condition (sclerosteosis) characterized by bone mineral contents which are one- to four-fold higher than in normal individuals. Thus, as discussed in more detail below this discovery has led to the development of assays which may be utilized to select molecules which inhibit the binding of the TGF-beta binding-protein to the TGF-beta family of proteins and bone morphogenic proteins (BMPs), and methods of utilizing such molecules for increasing the bone mineral content of warm-blooded animals (including for example, humans).

#### DISCUSSION OF THE DISEASE KNOWN AS SCLEROSTEOSIS

Sclerosteosis is a term that was applied by Hansen (1967) (Hansen, H. G., Sklerosteose. In: Opitz, H.; Schmid, F., Handbuch der Kinderheilkunde. Berlin: Springer (pub.) 6 1967. Pp. 351-355) to a disorder similar to van Buchem hyperostosis

corticalis generalisata but possibly differing in radiologic appearance of the bone changes and in the presence of asymmetric cutaneous syndactyly of the index and middle fingers in many cases.

Sclerosteosis is now known to be an autosomal semi-dominant disorder which is characterized by widely disseminated sclerotic lesions of the bone in the adult. The condition is progressive. Sclerosteosis also has a developmental aspect which is associated with syndactyly (two or more fingers are fused together). The Sclerosteosis Syndrome is associated with large stature and many affected individuals attain a height of six feet or more. The bone mineral content of homozygotes can be 1 to 6 fold over normal individuals and bone mineral density can be 1 to 4 fold above normal values (e.g., from unaffected siblings).

The Sclerosteosis Syndrome occurs primarily in Afrikaaners of Dutch descent in South Africa. Approximately 1/140 individuals in the Afrikaaner population are carriers of the mutated gene (heterozygotes). The mutation shows 100% penetrance. There are anecdotal reports of increased of bone mineral density in heterozygotes with no associated pathologies (syndactyly or skull overgrowth).

It appears at the present time that there is no abnormality of the pituitary-hypothalamus axis in Sclerosteosis. In particular, there appears to be no overproduction of growth hormone and cortisone. In addition, sex hormone levels are normal in affected individuals. However, bone turnover markers (osteoblast specific alkaline phosphatase, osteocalcin, type 1 procollagen C' propeptide (PICP), and total alkaline phosphatase; (see Comier, C., *Curr. Opin. in Rheu.* 7:243, 1995) indicate that there is hyperosteoblastic activity associated with the disease but that there is normal to slightly decreased osteoclast activity as measured by markers of bone resorption (pyridinoline, deoxypyridinoline, N-telopeptide, urinary hydroxyproline, plasma tartrate-resistant acid phosphatases and galactosyl hydroxylysine (see Comier, *supra*)).

Sclerosteosis is characterized by the continual deposition of bone throughout the skeleton during the lifetime of the affected individuals. In homozygotes the continual deposition of bone mineral leads to an overgrowth of bone in areas of the skeleton where there is an absence of mechanoreceptors (skull, jaw, cranium). In homozygotes with Sclerosteosis, the overgrowth of the bones of the skull leads to cranial compression and eventually to death due to excessive hydrostatic pressure on the brain stem. In all other parts of the skeleton there is a generalized and diffuse sclerosis. Cortical areas of the long bones are greatly thickened resulting in a substantial increase in bone strength. Trabecular connections are increased in thickness

which in turn increases the strength of the trabecular bone. Sclerotic bones appear unusually opaque to x-rays.

As described in more detail in Example 1, the rare genetic mutation that is responsible for the Sclerosteosis syndrome has been localized to the region of human chromosome 17 that encodes a novel member of the TGF-beta binding-protein family (one representative example of which is designated "H. Beer"). As described in more detail below, based upon this discovery, the mechanism of bone mineralization is more fully understood, allowing the development of assays for molecules which increase bone mineralization, and use of such molecules to increase bone mineral content, and in the treatment or prevention of a wide number of diseases.

#### TGF-BETA SUPER-FAMILY

The Transforming Growth Factor-beta (TGF-beta) super-family contains a variety of growth factors that share common sequence elements and structural motifs (at both the secondary and tertiary levels). This protein family is known to exert a wide spectrum of biological responses on a large variety of cell types. Many of them have important functions during the embryonal development in pattern formation and tissue specification; in adults they are involved, *e.g.*, in wound healing and bone repair and bone remodeling, and in the modulation of the immune system. In addition to the three TGF-beta's, the super-family includes the Bone Morphogenic Proteins (BMPs), Activins, Inhibins, Growth and Differentiation Factors (GDFs), and Glial-Derived Neurotrophic Factors (GDNFs). Primary classification is established through general sequence features that bin a specific protein into a general sub-family. Additional stratification within the sub-family is possible due to stricter sequence conservation between members of the smaller group. In certain instances, such as with BMP-5, BMP-6 and BMP-7, this can be as high as 75 percent amino acid homology between members of the smaller group. This level of identity enables a single representative sequence to illustrate the key biochemical elements of the sub-group that separates it from other members of the larger family.

TGF-beta signals by inducing the formation of hetero-oligomeric complexes of type I and type II receptors. The crystal structure of TGF-beta2 has been determined. The general fold of the TGF-beta2 monomer contains a stable, compact, cysteine knotlike structure formed by three disulphide bridges. Dimerization, stabilized by one disulphide bridge, is antiparallel.

TGF-beta family members initiate their cellular action by binding to receptors with intrinsic serine/threonine kinase activity. This receptor family consists

of two subfamilies, denoted type I and type II receptors. Each member of the TGF-beta family binds to a characteristic combination of type I and type II receptors, both of which are needed for signaling. In the current model for TGF-beta receptor activation, TGF-beta first binds to the type II receptor (TbR-II), which occurs in the cell  
5 membrane in an oligomeric form with activated kinase. Thereafter, the type I receptor (TbR-I), which can not bind ligand in the absence of TbR-II, is recruited into the complex. TbR-II then phosphorylates TbR-I predominantly in a domain rich in glycine and serine residues (GS domain) in the juxtamembrane region, and thereby activates TbR-I.

10 Thus far seven type I receptors and five type II receptors have been identified.

BONE MORPHOGENIC PROTEINS (BMPs) ARE KEY REGULATORY PROTEINS IN  
DETERMINING BONE MINERAL DENSITY IN HUMANS

A major advance in the understanding of bone formation was the  
15 identification of the bone morphogenic proteins (BMPs), also known as osteogenic proteins (OPs), which regulate cartilage and bone differentiation in vivo. BMPs/OPs induce endochondral bone differentiation through a cascade of events which include formation of cartilage, hypertrophy and calcification of the cartilage, vascular invasion, differentiation of osteoblasts, and formation of bone. As described above, the  
20 BMPs/OPs (BMP 2-14, and osteogenic protein 1 and -2, OP-1 and OP-2) are members of the TGF-beta super-family. The striking evolutionary conservation between members the BMP/OP sub-family suggests that they are critical in the normal development and function of animals. Moreover, the presence of multiple forms of BMPs/OPs raises an important question about the biological relevance of this apparent  
25 redundancy. In addition to postfetal chondrogenesis and osteogenesis, the BMPs/OPs play multiple roles in skeletogenesis (including the development of craniofacial and dental tissues) and in embryonic development and organogenesis of parenchymatous organs, including the kidney. It is now understood that nature relies on common (and few) molecular mechanisms tailored to provide the emergence of specialized tissues  
30 and organs. The BMP/OP super-family is an elegant example of nature parsimony in programming multiple specialized functions deploying molecular isoforms with minor variation in amino acid motifs within highly conserved carboxy-terminal regions.

BMP ANTAGONISM

The BMP and Activin sub-families are subject to significant post-



translational regulation. An intricate extracellular control system exists, whereby a high affinity antagonist is synthesized and exported, and subsequently complexes selectively with BMPs or activins to disrupt their biological activity (W.C. Smith (1999) *TIG* 15(1) 3-6). A number of these natural antagonists have been identified, and  
5 based on sequence divergence appear to have evolved independently due to the lack of primary sequence conservation. There has been no structural work to date on this class of proteins. Studies of these antagonists has highlighted a distinct preference for interacting and neutralizing BMP-2 and BMP-4. Furthermore, the mechanism of inhibition seems to differ for the different antagonists (S. Iemura et al. (1998) *Proc*  
10 *Natl Acad Sci USA* 95 9337-9342).

### NOVEL TGF-BETA BINDING-PROTEINS

#### 1. Background re: TGF-beta binding-proteins

As noted above, the present invention provides a novel class of TGF-beta binding-proteins that possess a nearly identical cysteine (disulfide) scaffold when  
15 compared to Human DAN, Human Gremlin, and Human Cerberus, and SCGF (U.S. Patent No. 5,780,263) but almost no homology at the nucleotide level (for background information, see generally Hsu, D.R., Economides, A.N., Wang, X., Eimon, P.M., Harland, R.M., "The *Xenopus* Dorsalizing Factor Gremlin Identifies a Novel Family of Secreted Proteins that Antagonize BMP Activities," *Molecular Cell* 1:673-683, 1998).

20 One representative example of the novel class of TGF-beta binding-proteins is disclosed in Sequence ID Nos. 1, 5, 9, 11, 13, and 15. Representative members of this class of binding proteins should also be understood to include variants of the TGF-beta binding-protein (e.g., Sequence ID Nos. 5 and 7). As utilized herein, a "TGF-beta binding-protein variant gene" refers to nucleic acid molecules that encode a  
25 polypeptide having an amino acid sequence that is a modification of SEQ ID Nos: 2, 10, 12, 14 or 16. Such variants include naturally-occurring polymorphisms or allelic variants of TGF-beta binding-protein genes, as well as synthetic genes that contain conservative amino acid substitutions of these amino acid sequences. Additional variant forms of a TGF-beta binding-protein gene are nucleic acid molecules that  
30 contain insertions or deletions of the nucleotide sequences described herein. TGF-beta binding-protein variant genes can be identified by determining whether the genes hybridize with a nucleic acid molecule having the nucleotide sequence of SEQ ID Nos: 1, 5, 7, 9, 11, 13, or 15 under stringent conditions. In addition, TGF-beta binding-protein variant genes should encode a protein having a cysteine backbone.

As an alternative, TGF-beta binding-protein variant genes can be identified by sequence comparison. As used herein, two amino acid sequences have "100% amino acid sequence identity" if the amino acid residues of the two amino acid sequences are the same when aligned for maximal correspondence. Similarly, two  
5 nucleotide sequences have "100% nucleotide sequence identity" if the nucleotide residues of the two nucleotide sequences are the same when aligned for maximal correspondence. Sequence comparisons can be performed using standard software programs such as those included in the LASERGENE bioinformatics computing suite, which is produced by DNASTAR (Madison, Wisconsin). Other methods for  
10 comparing two nucleotide or amino acid sequences by determining optimal alignment are well-known to those of skill in the art (see, for example, Peruski and Peruski, *The Internet and the New Biology: Tools for Genomic and Molecular Research* (ASM Press, Inc. 1997), Wu et al. (eds.), "Information Superhighway and Computer Databases of Nucleic Acids and Proteins," in *Methods in Gene Biotechnology*, pages  
15 123-151 (CRC Press, Inc. 1997), and Bishop (ed.), *Guide to Human Genome Computing*, 2nd Edition (Academic Press, Inc. 1998)).

A variant TGF-beta binding-protein should have at least a 50% amino acid sequence identity to SEQ ID NOs: 2, 6, 10, 12, 14 or 16 and preferably, greater than 60%, 65%, 70%, 75%, 80%, 85%, 90%, or 95% identity. Alternatively, TGF-beta  
20 binding-protein variants can be identified by having at least a 70% nucleotide sequence identity to SEQ ID NOs: 1, 5, 9, 11, 13 or 15. Moreover, the present invention contemplates TGF-beta binding-protein gene variants having greater than 75%, 80%, 85%, 90%, or 95% identity to SEQ ID NO:1. Regardless of the particular method used to identify a TGF-beta binding-protein variant gene or variant TGF-beta binding-  
25 protein, a variant TGF-beta binding-protein or a polypeptide encoded by a variant TGF-beta binding-protein gene can be functionally characterized by, for example, its ability to bind to and/or inhibit the signaling of a selected member of the TGF-beta family of proteins, or by its ability to bind specifically to an anti-TGF-beta binding-protein antibody.

30 The present invention includes functional fragments of TGF-beta binding-protein genes. Within the context of this invention, a "functional fragment" of a TGF-beta binding-protein gene refers to a nucleic acid molecule that encodes a portion of a TGF-beta binding-protein polypeptide which either (1) possesses the above-noted function activity, or (2) specifically binds with an anti-TGF-beta binding-  
35 protein antibody. For example, a functional fragment of a TGF-beta binding-protein gene described herein comprises a portion of the nucleotide sequence of SEQ ID Nos:

1. 5, 9, 11, 13, or 15.

2. Isolation of the TGF-beta binding-protein gene

DNA molecules encoding a binding-protein gene can be obtained by screening a human cDNA or genomic library using polynucleotide probes based upon, for example, SEQ ID NO:1.

For example, the first step in the preparation of a cDNA library is to isolate RNA using methods well-known to those of skill in the art. In general, RNA isolation techniques must provide a method for breaking cells, a means of inhibiting RNase-directed degradation of RNA, and a method of separating RNA from DNA, protein, and polysaccharide contaminants. For example, total RNA can be isolated by freezing tissue in liquid nitrogen, grinding the frozen tissue with a mortar and pestle to lyse the cells, extracting the ground tissue with a solution of phenol/chloroform to remove proteins, and separating RNA from the remaining impurities by selective precipitation with lithium chloride (see, for example, Ausubel et al. (eds.), *Short Protocols in Molecular Biology*, 3rd Edition, pages 4-1 to 4-6 (John Wiley & Sons 1995) ["Ausubel (1995)"]; Wu et al., *Methods in Gene Biotechnology*, pages 33-41 (CRC Press, Inc. 1997) ["Wu (1997)"]).

Alternatively, total RNA can be isolated by extracting ground tissue with guanidinium isothiocyanate, extracting with organic solvents, and separating RNA from contaminants using differential centrifugation (see, for example, Ausubel (1995) at pages 4-1 to 4-6; Wu (1997) at pages 33-41).

In order to construct a cDNA library, poly(A)<sup>+</sup> RNA must be isolated from a total RNA preparation. Poly(A)<sup>+</sup> RNA can be isolated from total RNA by using the standard technique of oligo(dT)-cellulose chromatography (see, for example, Ausubel (1995) at pages 4-11 to 4-12).

Double-stranded cDNA molecules are synthesized from poly(A)<sup>+</sup> RNA using techniques well-known to those in the art. (see, for example, Wu (1997) at pages 41-46). Moreover, commercially available kits can be used to synthesize double-stranded cDNA molecules. For example, such kits are available from Life Technologies, Inc. (Gaithersburg, Maryland), CLONTECH Laboratories, Inc. (Palo Alto, California), Promega Corporation (Madison, Wisconsin) and Stratagene Cloning Systems (La Jolla, California).

The basic approach for obtaining TGF-beta binding-protein cDNA clones can be modified by constructing a subtracted cDNA library which is enriched in TGF-binding-protein-specific cDNA molecules. Techniques for constructing subtracted libraries are well-known to those of skill in the art (see, for example, Sargent, "Isolation of

Differentially Expressed Genes," in *Meth. Enzymol.* 152:423, 1987, and Wu et al. (eds.), "Construction and Screening of Subtracted and Complete Expression cDNA Libraries," in *Methods in Gene Biotechnology*, pages 29-65 (CRC Press, Inc. 1997)).

Various cloning vectors are appropriate for the construction of a cDNA library. For example, a cDNA library can be prepared in a vector derived from bacteriophage, such as a  $\lambda$ gt10 vector (see, for example, Huynh et al., "Constructing and Screening cDNA Libraries in  $\lambda$ gt10 and  $\lambda$ gt11," in *DNA Cloning: A Practical Approach Vol. I*, Glover (ed.), page 49 (IRL Press, 1985); Wu (1997) at pages 47-52).

Alternatively, double-stranded cDNA molecules can be inserted into a plasmid vector, such as a pBluescript vector (Stratagene Cloning Systems; La Jolla, California), a LambdaGEM-4 (Promega Corp., Madison, Wisconsin) or other commercially available vectors. Suitable cloning vectors also can be obtained from the American Type Culture Collection (Rockville, Maryland).

In order to amplify the cloned cDNA molecules, the cDNA library is inserted into a prokaryotic host, using standard techniques. For example, a cDNA library can be introduced into competent *E. coli* DH5 cells, which can be obtained from Life Technologies, Inc. (Gaithersburg, Maryland).

A human genomic DNA library can be prepared by means well-known in the art (see, for example, Ausubel (1995) at pages 5-1 to 5-6; Wu (1997) at pages 307-327). Genomic DNA can be isolated by lysing tissue with the detergent Sarkosyl, digesting the lysate with proteinase K, clearing insoluble debris from the lysate by centrifugation, precipitating nucleic acid from the lysate using isopropanol, and purifying resuspended DNA on a cesium chloride density gradient.

DNA fragments that are suitable for the production of a genomic library can be obtained by the random shearing of genomic DNA or by the partial digestion of genomic DNA with restriction endonucleases. Genomic DNA fragments can be inserted into a vector, such as a bacteriophage or cosmid vector, in accordance with conventional techniques, such as the use of restriction enzyme digestion to provide appropriate termini, the use of alkaline phosphatase treatment to avoid undesirable joining of DNA molecules, and ligation with appropriate ligases. Techniques for such manipulation are well-known in the art (see, for example, Ausubel (1995) at pages 5-1 to 5-6; Wu (1997) at pages 307-327).

Nucleic acid molecules that encode a TGF-beta binding-protein gene can also be obtained using the polymerase chain reaction (PCR) with oligonucleotide primers having nucleotide sequences that are based upon the nucleotide sequences of the human TGF-beta binding-protein gene, as described herein. General methods for screening libraries with PCR are provided by, for example, Yu et al., "Use of the

Polymerase Chain Reaction to Screen Phage Libraries," in *Methods in Molecular Biology, Vol. 15: PCR Protocols: Current Methods and Applications*, White (ed.), pages 211-215 (Humana Press, Inc. 1993). Moreover, techniques for using PCR to isolate related genes are described by, for example, Preston, "Use of Degenerate  
5 Oligonucleotide Primers and the Polymerase Chain Reaction to Clone Gene Family Members," in *Methods in Molecular Biology, Vol. 15: PCR Protocols: Current Methods and Applications*, White (ed.), pages 317-337 (Humana Press, Inc. 1993).

Alternatively, human genomic libraries can be obtained from commercial sources such as Research Genetics (Huntsville, AL) and the American Type Culture  
10 Collection (Rockville, Maryland).

A library containing cDNA or genomic clones can be screened with one or more polynucleotide probes based upon SEQ ID NO:1, using standard methods (*see, for example*, Ausubel (1995) at pages 6-1 to 6-11).

Anti-TGF-beta binding-protein antibodies, produced as described below,  
15 can also be used to isolate DNA sequences that encode TGF-beta binding-protein genes from cDNA libraries. For example, the antibodies can be used to screen  $\lambda$ gt11 expression libraries, or the antibodies can be used for immunoscreening following hybrid selection and translation (*see, for example*, Ausubel (1995) at pages 6-12 to 6-16; Margolis et al., "Screening  $\lambda$  expression libraries with antibody and protein  
20 probes," in *DNA Cloning 2: Expression Systems, 2nd Edition*, Glover et al. (eds.), pages 1-14 (Oxford University Press 1995)).

The sequence of a TGF-beta binding-protein cDNA or TGF-beta binding-protein genomic fragment can be determined using standard methods. Moreover, the identification of genomic fragments containing a TGF-beta binding-  
25 protein promoter or regulatory element can be achieved using well-established techniques, such as deletion analysis (*see, generally*, Ausubel (1995)).

As an alternative, a TGF-beta binding-protein gene can be obtained by synthesizing DNA molecules using mutually priming long oligonucleotides and the nucleotide sequences described herein (*see, for example*, Ausubel (1995) at pages 8-8  
30 to 8-9). Established techniques using the polymerase chain reaction provide the ability to synthesize DNA molecules at least two kilobases in length (Adang et al., *Plant Molec. Biol.* 21:1131, 1993; Bambot et al., *PCR Methods and Applications* 2:266, 1993; Dillon et al., "Use of the Polymerase Chain Reaction for the Rapid Construction of Synthetic Genes," in *Methods in Molecular Biology, Vol. 15: PCR Protocols: Current Methods and Applications*, White (ed.), pages 263-268, (Humana Press, Inc.  
35 1993); Holowachuk et al., *PCR Methods Appl.* 4:299, 1995).

### 3. Production of TGF-beta binding-protein genes

Nucleic acid molecules encoding variant TGF-beta binding-protein genes can be obtained by screening various cDNA or genomic libraries with polynucleotide probes having nucleotide sequences based upon SEQ ID NO:1, 5, 9, 11, 13, or 15, using procedures described above. TGF-beta binding-protein gene variants can also be constructed synthetically. For example, a nucleic acid molecule can be devised that encodes a polypeptide having a conservative amino acid change, compared with the amino acid sequence of SEQ ID NOs: 2, 6, 8, 10, 12, 14, or 16. That is, variants can be obtained that contain one or more amino acid substitutions of SEQ ID NOs: 2, 6, 8, 10, 12, 14 or 16, in which an alkyl amino acid is substituted for an alkyl amino acid in a TGF-beta binding-protein amino acid sequence, an aromatic amino acid is substituted for an aromatic amino acid in a TGF-beta binding-protein amino acid sequence, a sulfur-containing amino acid is substituted for a sulfur-containing amino acid in a TGF-beta binding-protein amino acid sequence, a hydroxy-containing amino acid is substituted for a hydroxy-containing amino acid in a TGF-beta binding-protein amino acid sequence, an acidic amino acid is substituted for an acidic amino acid in a TGF-beta binding-protein amino acid sequence, a basic amino acid is substituted for a basic amino acid in a TGF-beta binding-protein amino acid sequence, or a dibasic monocarboxylic amino acid is substituted for a dibasic monocarboxylic amino acid in a TGF-beta binding-protein amino acid sequence.

Among the common amino acids, for example, a "conservative amino acid substitution" is illustrated by a substitution among amino acids within each of the following groups: (1) glycine, alanine, valine, leucine, and isoleucine, (2) phenylalanine, tyrosine, and tryptophan, (3) serine and threonine, (4) aspartate and glutamate, (5) glutamine and asparagine, and (6) lysine, arginine and histidine. In making such substitutions, it is important to, where possible, maintain the cysteine backbone outlined in Figure 1.

Conservative amino acid changes in a TGF-beta binding-protein gene can be introduced by substituting nucleotides for the nucleotides recited in SEQ ID NO:1. Such "conservative amino acid" variants can be obtained, for example, by oligonucleotide-directed mutagenesis, linker-scanning mutagenesis, mutagenesis using the polymerase chain reaction, and the like (see Ausubel (1995) at pages 8-10 to 8-22; and McPherson (ed.), *Directed Mutagenesis: A Practical Approach* (IRL Press 1991)). The functional ability of such variants can be determined using a standard method, such as the assay described herein. Alternatively, a variant TGF-beta binding-protein polypeptide can be identified by the ability to specifically bind anti-TGF-beta binding-

protein antibodies.

Routine deletion analyses of nucleic acid molecules can be performed to obtain "functional fragments" of a nucleic acid molecule that encodes a TGF-beta binding-protein polypeptide. As an illustration, DNA molecules having the nucleotide sequence of SEQ ID NO:1 can be digested with *Bal31* nuclease to obtain a series of nested deletions. The fragments are then inserted into expression vectors in proper reading frame, and the expressed polypeptides are isolated and tested for activity, or for the ability to bind anti-TGF-beta binding-protein antibodies. One alternative to exonuclease digestion is to use oligonucleotide-directed mutagenesis to introduce deletions or stop codons to specify production of a desired fragment. Alternatively, particular fragments of a TGF-beta binding-protein gene can be synthesized using the polymerase chain reaction.

Standard techniques for functional analysis of proteins are described by, for example, Treuter et al., *Molec. Gen. Genet.* 240:113, 1993; Content et al., "Expression and preliminary deletion analysis of the 42 kDa 2-5A synthetase induced by human interferon," in *Biological Interferon Systems, Proceedings of ISIR-TNO Meeting on Interferon Systems*, Cantell (ed.), pages 65-72 (Nijhoff 1987); Herschman, "The EGF Receptor," in *Control of Animal Cell Proliferation, Vol. 1*, Boynton et al., (eds.) pages 169-199 (Academic Press 1985); Coumailleau et al., *J. Biol. Chem.* 270:29270, 1995; Fukunaga et al., *J. Biol. Chem.* 270:25291, 1995; Yamaguchi et al., *Biochem. Pharmacol.* 50:1295, 1995; and Meisel et al., *Plant Molec. Biol.* 30:1, 1996.

The present invention also contemplates functional fragments of a TGF-beta binding-protein gene that have conservative amino acid changes.

A TGF-beta binding-protein variant gene can be identified on the basis of structure by determining the level of identity with nucleotide and amino acid sequences of SEQ ID NOs: 1, 5, 9, 11, 13, or 15 and 2, 6, 10, 12, 14, or 16, as discussed above. An alternative approach to identifying a variant gene on the basis of structure is to determine whether a nucleic acid molecule encoding a potential variant TGF-beta binding-protein gene can hybridize under stringent conditions to a nucleic acid molecule having the nucleotide sequence of SEQ ID Nos: 1, 5, 9, 11, 13, or 15, or a portion thereof of at least 15 or 20 nucleotides in length. As an illustration of stringent hybridization conditions, a nucleic acid molecule having a variant TGF-beta binding-protein sequence can bind with a fragment of a nucleic acid molecule having a sequence from SEQ ID NO:1 in a buffer containing, for example, 5xSSPE (1xSSPE = 180 mM sodium chloride, 10 mM sodium phosphate, 1 mM EDTA (pH 7.7), 5xDenhardt's solution (100xDenhardt's = 2% (w/v) bovine serum albumin, 2% (w/v)

Ficoll, 2% (w/v) polyvinylpyrrolidone) and 0.5% SDS incubated overnight at 55-60°C. Post-hybridization washes at high stringency are typically performed in 0.5xSSC (1xSSC = 150 mM sodium chloride, 15 mM trisodium citrate) or in 0.5xSSPE at 55-60°C.

5                   Regardless of the particular nucleotide sequence of a variant TGF-beta binding-protein gene, the gene encodes a polypeptide that can be characterized by its functional activity, or by the ability to bind specifically to an anti-TGF-beta binding-protein antibody. More specifically, variant TGF-beta binding-protein genes encode polypeptides which exhibit at least 50%, and preferably, greater than 60, 70, 80 or  
10                   90%, of the activity of polypeptides encoded by the human TGF-beta binding-protein gene described herein.

#### 4.           Production of TGF-beta binding-protein in Cultured Cells

To express a TGF-beta binding-protein gene, a nucleic acid molecule encoding the polypeptide must be operably linked to regulatory sequences that control  
15           transcriptional expression in an expression vector and then introduced into a host cell. In addition to transcriptional regulatory sequences, such as promoters and enhancers, expression vectors can include translational regulatory sequences and a marker gene which is suitable for selection of cells that carry the expression vector.

Expression vectors that are suitable for production of a foreign protein in  
20           eukaryotic cells typically contain (1) prokaryotic DNA elements coding for a bacterial replication origin and an antibiotic resistance marker to provide for the growth and selection of the expression vector in a bacterial host; (2) eukaryotic DNA elements that control initiation of transcription, such as a promoter; and (3) DNA elements that control the processing of transcripts, such as a transcription termination/polyadenylation sequence.

25           TGF-beta binding-proteins of the present invention are preferably expressed in mammalian cells. Examples of mammalian host cells include African green monkey kidney cells (Vero; ATCC CRL 1587), human embryonic kidney cells (293-HEK; ATCC CRL 1573), baby hamster kidney cells (BHK-21; ATCC CRL 8544), canine kidney cells (MDCK; ATCC CCL 34), Chinese hamster ovary cells (CHO-K1; ATCC CCL61), rat  
30           pituitary cells (GH1; ATCC CCL82), HeLa S3 cells (ATCC CCL2.2), rat hepatoma cells (H-4-II-E; ATCC CRL 1548) SV40-transformed monkey kidney cells (COS-1; ATCC CRL 1650) and murine embryonic cells (NIH-3T3; ATCC CRL 1658).

For a mammalian host, the transcriptional and translational regulatory signals may be derived from viral sources, such as adenovirus, bovine papilloma virus,  
35           simian virus, or the like, in which the regulatory signals are associated with a particular gene



which has a high level of expression. Suitable transcriptional and translational regulatory sequences also can be obtained from mammalian genes, such as actin, collagen, myosin, and metallothionein genes.

Transcriptional regulatory sequences include a promoter region sufficient to  
 5 direct the initiation of RNA synthesis. Suitable eukaryotic promoters include the promoter of the mouse metallothionein I gene [Hamer et al., *J. Molec. Appl. Genet.* 1:273, 1982], the TK promoter of Herpes virus [McKnight, *Cell* 31:355, 1982], the SI'40 early promoter [Benoist et al., *Nature* 290:304, 1981], the Rous sarcoma virus promoter [Gorman et al.,  
 10 *Proc. Nat'l Acad. Sci. USA* 79:6777, 1982], the cytomegalovirus promoter [Foecking et al., *Gene* 45:101, 1980], and the mouse mammary tumor virus promoter (see, generally, Etcheverry, "Expression of Engineered Proteins in Mammalian Cell Culture," in *Protein Engineering: Principles and Practice*, Cleland et al. (eds.), pages 163-181 (John Wiley & Sons, Inc. 1996)).

Alternatively, a prokaryotic promoter, such as the bacteriophage T3 RNA  
 15 polymerase promoter, can be used to control TGF-beta binding-protein gene expression in mammalian cells if the prokaryotic promoter is regulated by a eukaryotic promoter (Zhou et al., *Mol. Cell. Biol.* 10:4529, 1990; Kaufman et al., *Nucl. Acids Res.* 19:4485, 1991).

TGF-beta binding-protein genes may also be expressed in bacterial, yeast, insect, or plant cells. Suitable promoters that can be used to express TGF-beta binding-  
 20 protein polypeptides in a prokaryotic host are well-known to those of skill in the art and include promoters capable of recognizing the T4, T3, Sp6 and T7 polymerases, the P<sub>R</sub> and P<sub>L</sub> promoters of bacteriophage lambda, the *trp*, *recA*, heat shock, *lacUV5*, *tac*, *lpp-lacSpr*, *phoA*, and *lacZ* promoters of *E. coli*, promoters of *B. subtilis*, the promoters of the bacteriophages of *Bacillus*, *Streptomyces* promoters, the *int* promoter of bacterio-  
 25 phage lambda, the *hla* promoter of pBR322, and the CAT promoter of the chloramphenicol acetyl transferase gene. Prokaryotic promoters have been reviewed by Glick, *J. Ind. Microbiol.* 1:277, 1987, Watson et al., *Molecular Biology of the Gene*, 4th Ed. (Benjamin Cummins 1987), and by Ausubel et al. (1995).

Preferred prokaryotic hosts include *E. coli* and *Bacillus subtilis*.  
 30 Suitable strains of *E. coli* include BL21(DE3), BL21(DE3)pLysS, BL21(DE3)pLysE, DH1, DH41, DH5, DH51, DH5IF', DH5IMCR, DH10B, DH10B/p3, DH11S, C600, HB101, JM101, JM105, JM109, JM110, K38, RR1, Y1088, Y1089, CSH18, ER1451, and ER1647 (see, for example, Brown (Ed.), *Molecular Biology Labfax* (Academic Press 1991)). Suitable strains of *Bacillus subtilis* include BR151, YB886, MI119,  
 35 MI120, and B170 (see, for example, Hardy, "Bacillus Cloning Methods," in *DNA Cloning: A Practical Approach*, Glover (Ed.) (IRL Press 1985)).

Methods for expressing proteins in prokaryotic hosts are well-known to those of skill in the art (see, for example, Williams et al., "Expression of foreign proteins in *E. coli* using plasmid vectors and purification of specific polyclonal antibodies," in *DNA Cloning 2: Expression Systems, 2nd Edition*, Glover et al. (eds.), page 15 (Oxford University Press 1995); Ward et al., "Genetic Manipulation and Expression of Antibodies," in *Monoclonal Antibodies: Principles and Applications*, page 137 (Wiley-Liss, Inc. 1995); and Georgiou, "Expression of Proteins in Bacteria," in *Protein Engineering: Principles and Practice*, Cleland et al. (eds.), page 101 (John Wiley & Sons, Inc. 1996)).

The baculovirus system provides an efficient means to introduce cloned *TGF-beta binding-protein* genes into insect cells. Suitable expression vectors are based upon the *Autographa californica* multiple nuclear polyhedrosis virus (AcMNPV), and contain well-known promoters such as *Drosophila* heat shock protein (hsp) 70 promoter, *Autographa californica* nuclear polyhedrosis virus immediate-early gene promoter (*ie-1*) and the delayed early 39K promoter, baculovirus p10 promoter, and the *Drosophila* metallothionein promoter. Suitable insect host cells include cell lines derived from IPLB-Sf-21, a *Spodoptera frugiperda* pupal ovarian cell line, such as Sf9 (ATCC CRL 1711), Sf21AE, and Sf21 (Invitrogen Corporation; San Diego, CA), as well as *Drosophila* Schneider-2 cells. Established techniques for producing recombinant proteins in baculovirus systems are provided by Bailey et al., "Manipulation of Baculovirus Vectors," in *Methods in Molecular Biology, Volume 7: Gene Transfer and Expression Protocols*, Murray (ed.), pages 147-168 (The Humana Press, Inc. 1991), by Patel et al., "The baculovirus expression system," in *DNA Cloning 2: Expression Systems, 2nd Edition*, Glover et al. (eds.), pages 205-244 (Oxford University Press 1995), by Ausubel (1995) at pages 16-37 to 16-57, by Richardson (ed.), *Baculovirus Expression Protocols* (The Humana Press, Inc. 1995), and by Lucknow, "Insect Cell Expression Technology," in *Protein Engineering: Principles and Practice*, Cleland et al. (eds.), pages 183-218 (John Wiley & Sons, Inc. 1996).

Promoters for expression in yeast include promoters from *GAL1* (galactose), *PGK* (phosphoglycerate kinase), *ADH* (alcohol dehydrogenase), *AOX1* (alcohol oxidase), *HIS4* (histidinol dehydrogenase), and the like. Many yeast cloning vectors have been designed and are readily available. These vectors include YIp-based vectors, such as YIp5, YRp vectors, such as YRp17, YEpl vectors such as YEpl3 and YCp vectors, such as YCp19. One skilled in the art will appreciate that there are a wide variety of suitable vectors for expression in yeast cells.

Expression vectors can also be introduced into plant protoplasts, intact plant

tissues, or isolated plant cells. General methods of culturing plant tissues are provided, for example, by Miki et al., "Procedures for Introducing Foreign DNA into Plants," in *Methods in Plant Molecular Biology and Biotechnology*, Glick et al. (eds.), pages 67-88 (CRC Press, 1993).

5 An expression vector can be introduced into host cells using a variety of standard techniques including calcium phosphate transfection, liposome-mediated transfection, microprojectile-mediated delivery, electroporation, and the like. Preferably, the transfected cells are selected and propagated to provide recombinant host cells that comprise the expression vector stably integrated in the host cell genome. Techniques for  
10 introducing vectors into eukaryotic cells and techniques for selecting such stable transformants using a dominant selectable marker are described, for example, by Ausubel (1995) and by Murray (ed.), *Gene Transfer and Expression Protocols* (Humana Press 1991). Methods for introducing expression vectors into bacterial, yeast, insect, and plant cells are also provided by Ausubel (1995).

15 General methods for expressing and recovering foreign protein produced by a mammalian cell system is provided by, for example, Etcheverry, "Expression of Engineered Proteins in Mammalian Cell Culture," in *Protein Engineering: Principles and Practice*, Cleland et al. (eds.), pages 163 (Wiley-Liss, Inc. 1996). Standard techniques for recovering protein produced by a bacterial system is provided by, for example,  
20 Grisshammer et al., "Purification of over-produced proteins from *E. coli* cells," in *DNA Cloning 2: Expression Systems, 2nd Edition*, Glover et al. (eds.), pages 59-92 (Oxford University Press 1995). Established methods for isolating recombinant proteins from a baculovirus system are described by Richardson (ed.), *Baculovirus Expression Protocols* (The Humana Press, Inc., 1995).

25 More generally, TGF-beta binding-protein can be isolated by standard techniques, such as affinity chromatography, size exclusion chromatography, ion exchange chromatography, HPLC and the like. Additional variations in TGF-beta binding-protein isolation and purification can be devised by those of skill in the art. For example, anti-TGF-beta binding-protein antibodies, obtained as described below,  
30 can be used to isolate large quantities of protein by immunoaffinity purification.

#### 5. Production of Antibodies to TGF-beta binding-proteins

Antibodies to TGF-beta binding-protein can be obtained, for example, using the product of an expression vector as an antigen. Particularly useful anti-TGF-beta binding-protein antibodies "bind specifically" with TGF-beta binding-protein of  
35 Sequence ID Nos. 2, 6, 10, 12, 14, or 16, but not to other TGF-beta binding-proteins

such as Dan, Cerberus, SCGF, or Gremlin. Antibodies of the present invention (including fragments and derivatives thereof) may be a polyclonal or, especially a monoclonal antibody. The antibody may belong to any immunoglobulin class, and may be for example an IgG, for example IgG<sub>1</sub>, IgG<sub>2</sub>, IgG<sub>3</sub>, IgG<sub>4</sub>, IgE, IgM, or IgA antibody. It may be of animal, for example mammalian origin, and may be for example a murine, rat, human or other primate antibody. Where desired the antibody may be an internalising antibody.

Polyclonal antibodies to recombinant TGF-beta binding-protein can be prepared using methods well-known to those of skill in the art (see, for example, Green et al., "Production of Polyclonal Antisera," in *Immunochemical Protocols* (Manson, ed.), pages 1-5 (Humana Press 1992); Williams et al., "Expression of foreign proteins in *E. coli* using plasmid vectors and purification of specific polyclonal antibodies," in *DNA Cloning 2: Expression Systems, 2nd Edition*, Glover et al. (eds.), page 15 (Oxford University Press 1995)). Although polyclonal antibodies are typically raised in animals such as rats, mice, rabbits, goats, or sheep, an anti-TGF-beta binding-protein antibody of the present invention may also be derived from a subhuman primate antibody. General techniques for raising diagnostically and therapeutically useful antibodies in baboons may be found, for example, in Goldenberg et al., international patent publication No. WO 91/11465 (1991), and in Losman et al., *Int. J. Cancer* 46:310, 1990.

The antibody should comprise at least a variable region domain. The variable region domain may be of any size or amino acid composition and will generally comprise at least one hypervariable amino acid sequence responsible for antigen binding embedded in a framework sequence. In general terms the variable (V) region domain may be any suitable arrangement of immunoglobulin heavy (V<sub>H</sub>) and/or light (V<sub>L</sub>) chain variable domains. Thus for example the V region domain may be monomeric and be a V<sub>H</sub> or V<sub>L</sub> domain where these are capable of independently binding antigen with acceptable affinity. Alternatively the V region domain may be dimeric and contain V<sub>H</sub>-V<sub>H</sub>, V<sub>H</sub>-V<sub>L</sub>, or V<sub>L</sub>-V<sub>L</sub>, dimers in which the V<sub>H</sub> and V<sub>L</sub> chains are non-covalently associated (abbreviated hereinafter as F<sub>v</sub>). Where desired, however, the chains may be covalently coupled either directly, for example via a disulphide bond between the two variable domains, or through a linker, for example a peptide linker, to form a single chain domain (abbreviated hereinafter as scF<sub>v</sub>).

The variable region domain may be any naturally occurring variable domain or an engineered version thereof. By engineered version is meant a variable region domain which has been created using recombinant DNA engineering techniques. Such engineered

versions include those created for example from natural antibody variable regions by insertions, deletions or changes in or to the amino acid sequences of the natural antibodies. Particular examples of this type include those engineered variable region domains containing at least one CDR and optionally one or more framework amino acids from one antibody and the remainder of the variable region domain from a second antibody.

The variable region domain may be covalently attached at a C-terminal amino acid to at least one other antibody domain or a fragment thereof. Thus, for example where a  $V_H$  domain is present in the variable region domain this may be linked to an immunoglobulin  $C_{H1}$  domain or a fragment thereof. Similarly a  $V_L$  domain may be linked to a  $C_K$  domain or a fragment thereof. In this way for example the antibody may be a Fab fragment wherein the antigen binding domain contains associated  $V_H$  and  $V_L$  domains covalently linked at their C-termini to a  $CH1$  and  $C_K$  domain respectively. The  $CH1$  domain may be extended with further amino acids, for example to provide a hinge region domain as found in a Fab' fragment, or to provide further domains, such as antibody  $CH2$  and  $CH3$  domains.

Another form of an antibody fragment is a peptide coding for a single complementarity-determining region (CDR). CDR peptides ("minimal recognition units") can be obtained by constructing genes encoding the CDR of an antibody of interest. Such genes are prepared, for example, by using the polymerase chain reaction to synthesize the variable region from RNA of antibody-producing cells (see, for example, Larrick et al., *Methods: A Companion to Methods in Enzymology* 2:106, 1991; Courtenay-Luck, "Genetic Manipulation of Monoclonal Antibodies," in *Monoclonal Antibodies: Production, Engineering and Clinical Application*, Ritter et al. (eds.), page 166 (Cambridge University Press 1995); and Ward et al., "Genetic Manipulation and Expression of Antibodies," in *Monoclonal Antibodies: Principles and Applications*, Birch et al., (eds.), page 137 (Wiley-Liss, Inc. 1995)).

Antibodies for use in the invention may in general be monoclonal (prepared by conventional immunisation and cell fusion procedures) or in the case of fragments, derived therefrom using any suitable standard chemical e.g. reduction or enzymatic cleavage and/or digestion techniques, for example by treatment with pepsin.

More specifically, monoclonal anti-TGF-beta binding-protein antibodies can be generated utilizing a variety of techniques. Rodent monoclonal antibodies to specific antigens may be obtained by methods known to those skilled in the art (see, for example, Kohler et al., *Nature* 256:495, 1975; and Coligan et al. (eds.), *Current Protocols in Immunology*, 1:2.5.1-2.6.7 (John Wiley & Sons 1991) ["Coligan"]; Picksley et al., "Production of monoclonal antibodies against proteins expressed in *E.*

*coli*," in *DNA Cloning 2: Expression Systems, 2nd Edition*, Glover et al. (eds.), page 93 (Oxford University Press 1995)).

Briefly, monoclonal antibodies can be obtained by injecting mice with a composition comprising a TGF-beta binding-protein gene product, verifying the presence of antibody production by removing a serum sample, removing the spleen to obtain B-lymphocytes, fusing the B-lymphocytes with myeloma cells to produce hybridomas, cloning the hybridomas, selecting positive clones which produce antibodies to the antigen, culturing the clones that produce antibodies to the antigen, and isolating the antibodies from the hybridoma cultures.

In addition, an anti-TGF-beta binding-protein antibody of the present invention may be derived from a human monoclonal antibody. Human monoclonal antibodies are obtained from transgenic mice that have been engineered to produce specific human antibodies in response to antigenic challenge. In this technique, elements of the human heavy and light chain locus are introduced into strains of mice derived from embryonic stem cell lines that contain targeted disruptions of the endogenous heavy chain and light chain loci. The transgenic mice can synthesize human antibodies specific for human antigens, and the mice can be used to produce human antibody-secreting hybridomas. Methods for obtaining human antibodies from transgenic mice are described, for example, by Green et al., *Nature Genet.* 7:13, 1994; Lonberg et al., *Nature* 368:856, 1994; and Taylor et al., *Int. Immun.* 6:579, 1994.

Monoclonal antibodies can be isolated and purified from hybridoma cultures by a variety of well-established techniques. Such isolation techniques include affinity chromatography with Protein-A Sepharose, size-exclusion chromatography, and ion-exchange chromatography (see, for example, Coligan at pages 2.7.1-2.7.12 and pages 2.9.1-2.9.3; Baines et al., "Purification of Immunoglobulin G (IgG)," in *Methods in Molecular Biology, Vol. 10*, pages 79-104 (The Humana Press, Inc. 1992)).

For particular uses, it may be desirable to prepare fragments of anti-TGF-beta binding-protein antibodies. Such antibody fragments can be obtained, for example, by proteolytic hydrolysis of the antibody. Antibody fragments can be obtained by pepsin or papain digestion of whole antibodies by conventional methods. As an illustration, antibody fragments can be produced by enzymatic cleavage of antibodies with pepsin to provide a 5S fragment denoted  $F(ab')_2$ . This fragment can be further cleaved using a thiol reducing agent to produce 3.5S Fab' monovalent fragments. Optionally, the cleavage reaction can be performed using a blocking group for the sulfhydryl groups that result from cleavage of disulfide linkages. As an alternative, an enzymatic cleavage using pepsin produces two monovalent Fab

fragments and an Fc fragment directly. These methods are described, for example, by Goldenberg, U.S. patent No. 4,331,647, Nisonoff et al., *Arch Biochem. Biophys.* 89:230, 1960, Porter, *Biochem. J.* 73:119, 1959, Edelman et al., in *Methods in Enzymology* 1:422 (Academic Press 1967), and by Coligan at pages 2.8.1-2.8.10 and 2.10.-2.10.4.

Other methods of cleaving antibodies, such as separation of heavy chains to form monovalent light-heavy chain fragments, further cleavage of fragments, or other enzymatic, chemical or genetic techniques may also be used, so long as the fragments bind to the antigen that is recognized by the intact antibody.

Alternatively, the antibody may be a recombinant or engineered antibody obtained by the use of recombinant DNA techniques involving the manipulation and re-expression of DNA encoding antibody variable and/or constant regions. Such DNA is known and/or is readily available from DNA libraries including for example phage-antibody libraries (see Chiswell, D J and McCafferty, J. *Tibtech.* 10 80-84 (1992)) or where desired can be synthesised. Standard molecular biology and/or chemistry procedures may be used to sequence and manipulate the DNA, for example, to introduce codons to create cysteine residues, to modify, add or delete other amino acids or domains as desired.

From here, one or more replicable expression vectors containing the DNA may be prepared and used to transform an appropriate cell line, e.g. a non-producing myeloma cell line, such as a mouse NSO line or a bacterial, e.g. *E.coli* line, in which production of the antibody will occur. In order to obtain efficient transcription and translation, the DNA sequence in each vector should include appropriate regulatory sequences, particularly a promoter and leader sequence operably linked to the variable domain sequence. Particular methods for producing antibodies in this way are generally well known and routinely used. For example, basic molecular biology procedures are described by Maniatis *et al* (Molecular Cloning, Cold Spring Harbor Laboratory, New York, 1989); DNA sequencing can be performed as described in Sanger *et al* (PNAS 74, 5463, (1977)) and the Amersham International plc sequencing handbook; and site directed mutagenesis can be carried out according to the method of Kramer *et al* (Nucl. Acids Res. 12, 9441, (1984)) and the Anglian Biotechnology Ltd handbook. Additionally, there are numerous publications, detailing techniques suitable for the preparation of antibodies by manipulation of DNA, creation of expression vectors and transformation of appropriate cells, for example as reviewed by Mountain A and Adair, J R in *Biotechnology and Genetic Engineering Reviews* (ed. Tombs, M P, 10, Chapter 1, 1992, Intercept, Andover, UK) and in International Patent Specification No. WO 91/09967.

Where desired, the antibody according to the invention may have one or

more effector or reporter molecules attached to it and the invention extends to such modified proteins. The effector or reporter molecules may be attached to the antibody through any available amino acid side-chain, terminal amino acid or, where present carbohydrate functional group located in the antibody, always provided of course that this  
5 does not adversely affect the binding properties and eventual usefulness of the molecule. Particular functional groups include, for example any free amino, imino, thiol, hydroxyl, carboxyl or aldehyde group. Attachment of the antibody and the effector and/or reporter molecule(s) may be achieved via such groups and an appropriate functional group in the effector or reporter molecules. The linkage may be direct or indirect, through spacing or  
10 bridging groups.

Effector molecules include, for example, antineoplastic agents, toxins (such as enzymatically active toxins of bacterial or plant origin and fragments thereof e.g. ricin and fragments thereof) biologically active proteins, for example enzymes, nucleic acids and fragments thereof, e.g. DNA, RNA and fragments thereof, naturally occurring and synthetic  
15 polymers e.g. polysaccharides and polyalkylene polymers such as poly(ethylene glycol) and derivatives thereof, radionuclides, particularly radioiodide, and chelated metals. Suitable reporter groups include chelated metals, fluorescent compounds or compounds which may be detected by NMR or ESR spectroscopy.

Particular antineoplastic agents include cytotoxic and cytostatic agents, for  
20 example alkylating agents, such as nitrogen mustards (e.g. chlorambucil, melphalan, mechlorethamine, cyclophosphamide, or uracil mustard) and derivatives thereof, triethylenephosphoramidate, triethylenethiophosphor-amide, busulphan, or cisplatin; antimetabolites, such as methotrexate, fluorouracil, floxuridine, cytarabine, mercaptopurine, thioguanine, fluoroacetic acid or fluorocitric acid, antibiotics, such as bleomycins (e.g.  
25 bleomycin sulphate), doxorubicin, daunorubicin, mitomycins (e.g. mitomycin C), actinomycins (e.g. dactinomycin) plicamycin, calichaemicin and derivatives thereof, or esperamicin and derivatives thereof; mitotic inhibitors, such as etoposide, vincristine or vinblastine and derivatives thereof; alkaloids, such as ellipticine; polyols such as taxicin-I or taxicin-II; hormones, such as androgens (e.g. dromostanolone or testolactone), progestins  
30 (e.g. megestrol acetate or medroxyprogesterone acetate), estrogens (e.g. dimethylstilbestrol diphosphate, polyestradiol phosphate or estramustine phosphate) or antiestrogens (e.g. tamoxifen); anthraquinones, such as mitoxantrone, ureas, such as hydroxyurea; hydrazines, such as procarbazine; or imidazoles, such as dacarbazine.

Particularly useful effector groups are calichaemicin and derivatives thereof  
35 (see for example South African Patent Specifications Nos. 85/8794, 88/8127 and 90/2839).

Chelated metals include chelates of di-or tripositive metals having a



coordination number from 2 to 8 inclusive. Particular examples of such metals include technetium (Tc), rhenium (Re), cobalt (Co), copper (Cu), gold (Au), silver (Ag), lead (Pb), bismuth (Bi), indium (In), gallium (Ga), yttrium (Y), terbium (Tb), gadolinium (Gd), and scandium (Sc). In general the metal is preferably a radionuclide. Particular radionuclides include <sup>99m</sup>Tc, <sup>186</sup>Re, <sup>188</sup>Re, <sup>58</sup>Co, <sup>60</sup>Co, <sup>67</sup>Cu, <sup>195</sup>Au, <sup>199</sup>Au, <sup>110</sup>Ag, <sup>203</sup>Pb, <sup>206</sup>Bi, <sup>207</sup>Bi, <sup>111</sup>In, <sup>67</sup>Ga, <sup>68</sup>Ga, <sup>88</sup>Y, <sup>90</sup>Y, <sup>160</sup>Tb, <sup>153</sup>Gd and <sup>47</sup>Sc.

The chelated metal may be for example one of the above types of metal chelated with any suitable polydentate chelating agent, for example acyclic or cyclic polyamines, polyethers, (e.g. crown ethers and derivatives thereof); polyamides; porphyrins; and carbocyclic derivatives.

In general, the type of chelating agent will depend on the metal in use. One particularly useful group of chelating agents in conjugates according to the invention, however, are acyclic and cyclic polyamines, especially polyaminocarboxylic acids, for example diethylenetriaminepentaacetic acid and derivatives thereof, and macrocyclic amines, e.g. cyclic tri-aza and tetra-aza derivatives (for example as described in International Patent Specification No. WO 92/22583); and polyamides, especially desferrioxamine and derivatives thereof.

Thus for example when it is desired to use a thiol group in the antibody as the point of attachment this may be achieved through reaction with a thiol reactive group present in the effector or reporter molecule. Examples of such groups include an  $\alpha$ -halocarboxylic acid or ester, e.g. iodoacetamide, an imide, e.g. maleimide, a vinyl sulphone, or a disulphide. These and other suitable linking procedures are generally and more particularly described in International Patent Specifications Nos. WO 93/06231, WO 92/22583, WO 90/091195 and WO 89/01476.

## 25 ASSAYS FOR SELECTING MOLECULES WHICH INCREASE BONE DENSITY

As discussed above, the present invention provides methods for selecting and/or isolating compounds which are capable of increasing bone density. For example, within one aspect of the present invention methods are provided for determining whether a selected molecule is capable of increasing bone mineral content, comprising the steps of (a) mixing a selected molecule with TGF-beta binding protein and a selected member of the TGF-beta family of proteins, (b) determining whether the selected molecule stimulates signaling by the TGF-beta family of proteins, or inhibits the binding of the TGF-beta binding protein to the TGF-beta family of proteins. Within certain embodiments, the molecule enhances the ability of TGF-beta to function as a positive regulator of mesenchymal cell differentiation.

Within other aspects of the invention, methods are provided for determining whether a selected molecule is capable of increasing bone mineral content, comprising the steps of (a) exposing a selected molecule to cells which express TGF-beta binding-protein and (b) determining whether the expression (or activity) of TGF-beta binding-protein from said exposed cells decreases, and therefrom determining whether the compound is capable of increasing bone mineral content. Within one embodiment, the cells are selected from the group consisting of the spontaneously transformed or untransformed normal human bone from bone biopsies and rat parietal bone osteoblasts. Such methods may be accomplished in a wide variety of assay formats including, for example, Countercurrent Immuno-Electrophoresis (CIEP), Radioimmunoassays, Radioimmunoprecipitations, Enzyme-Linked Immuno-Sorbent Assays (ELISA), Dot Blot assays, Inhibition or Competition assays, and sandwich assays (*see* U.S. Patent Nos. 4,376,110 and 4,486,530; *see also Antibodies: A Laboratory Manual, supra*).

Representative embodiments of such assays are provided below in Examples 5 and 6. Briefly, a family member of the TGF-beta super-family or a TGF-beta binding protein is first bound to a solid phase, followed by addition of a candidate molecule. The labeled family member of the TGF-beta super-family or a TGF-beta binding protein is then added to the assay, the solid phase washed, and the quantity of bound or labeled TGF-beta super-family member or TGF-beta binding protein on the solid support determined. Molecules which are suitable for use in increasing bone mineral content as described herein are those molecules which decrease the binding of TGF-beta binding protein to a member or members of the TGF-beta super-family in a statistically significant manner. Obviously, assays suitable for use within the present invention should not be limited to the embodiments described within Examples 2 and 3. In particular, numerous parameters may be altered, such as by binding TGF-beta to a solid phase, or by elimination of a solid phase entirely.

Within other aspects of the invention, methods are provided for determining whether a selected molecule is capable of increasing bone mineral content, comprising the steps of (a) exposing a selected molecule to cells which express TGF-beta and (b) determining whether the activity of TGF-beta from said exposed cells is altered, and therefrom determining whether the compound is capable of increasing bone mineral content. Similar to the above described methods, a wide variety of methods may be utilized to assess the changes of TGF-beta binding-protein expression due to a selected test compound.

For example, within one aspect of the present invention methods are provided for determining whether a selected molecule is capable of increasing bone mineral content, comprising the steps of (a) mixing a selected molecule with TGF-beta-binding-protein and a selected member of the TGF-beta family of proteins, (b) determining whether the selected molecule up-regulates the signaling of the TGF-beta family of proteins, or inhibits the binding of the TGF-beta binding-protein to the TGF-beta family of proteins. Within certain embodiments, the molecule enhances the ability of TGF-beta to function as a positive regulator of mechemchymal cell differentiation.

Similar to the above described methods, a wide variety of methods may be utilized to assess stimulation of TGF-beta due to a selected test compound. One such representative method is provided below in Example 6 (see also Durham et al., *Endo.* 136:1374-1380).

Within yet other aspects of the present invention, methods are provided for determining whether a selected molecule is capable of increasing bone mineral content, comprising the step of determining whether a selected molecule inhibits the binding of TGF-beta binding-protein to bone, or an analogue thereof. As utilized herein, it should be understood that bone or analogues thereof refers to hydroxyapatite, or a surface composed of a powdered form of bone, crushed bone or intact bone. Similar to the above described methods, a wide variety of methods may be utilized to assess the inhibition of TGF-beta binding-protein localization to bone matrix. One such representative method is provided below in Example 7.

It should be noted that while the methods recited herein may refer to the analysis of an individual test molecule, that the present invention should not be so limited. In particular, the selected molecule may be contained within a mixture of compounds. Hence, the recited methods may further comprise the step of isolating a molecule which inhibits the binding of TGF-beta binding-protein to a TGF-beta family member.

#### CANDIDATE MOLECULES

A wide variety of molecules may be assayed for their ability to inhibit the binding of TGF-beta binding-protein to a TGF-beta family member. Representative examples which are discussed in more detail below include organic molecules, proteins or peptides, and nucleic acid molecules. Although it should be evident from the discussion below that the candidate molecules described herein may be utilized in the

assays described herein, it should also be readily apparent that such molecules can also be utilized in a variety of diagnostic and therapeutic settings.

### 1. Organic Molecules

Numerous organic molecules may be assayed for their ability to inhibit the binding of TGF-beta binding-protein to a TGF-beta family member.

For example, within one embodiment of the invention suitable organic molecules may be selected from either a chemical library, wherein chemicals are assayed individually, or from combinatorial chemical libraries where multiple compounds are assayed at once, then deconvoluted to determine and isolate the most active compounds.

Representative examples of such combinatorial chemical libraries include those described by Agrafiotis et al., "System and method of automatically generating chemical compounds with desired properties," U.S. Patent No. 5,463,564; Armstrong, R.W., "Synthesis of combinatorial arrays of organic compounds through the use of multiple component combinatorial array syntheses," WO 95/02566; Baldwin, J.J. et al., "Sulfonamide derivatives and their use," WO 95/24186; Baldwin, J.J. et al., "Combinatorial dihydrobenzopyran library," WO 95/30642; Brenner, S., "New kit for preparing combinatorial libraries," WO 95/16918; Chenera, B. et al., "Preparation of library of resin-bound aromatic carbocyclic compounds," WO 95/16712; Ellman, J.A., "Solid phase and combinatorial synthesis of benzodiazepine compounds on a solid support," U.S. Patent No. 5,288,514; Felder, E. et al., "Novel combinatorial compound libraries," WO 95/16209; Lerner, R. et al., "Encoded combinatorial chemical libraries," WO 93/20242; Pavia, M.R. et al., "A method for preparing and selecting pharmaceutically useful non-peptide compounds from a structurally diverse universal library," WO 95/04277; Summerton, J.E. and D.D. Weller, "Morpholino-subunit combinatorial library and method," U.S. Patent No. 5,506,337; Holmes, C., "Methods for the Solid Phase Synthesis of Thiazolidinones, Metathiazanones, and Derivatives thereof," WO 96/00148; Phillips, G.B. and G.P. Wei, "Solid-phase Synthesis of Benzimidazoles," *Tet. Letters* 37:4887-90, 1996; Ruhland, B. et al., "Solid-supported Combinatorial Synthesis of Structurally Diverse  $\beta$ -Lactams," *J. Amer. Chem. Soc.* 118:253-4, 1996; Look, G.C. et al., "The Identification of Cyclooxygenase-1 Inhibitors from 4-Thiazolidinone Combinatorial Libraries," *Bioorg and Med. Chem. Letters* 6:707-12, 1996.

## 2. Proteins and Peptides

A wide range of proteins and peptides may likewise be utilized as candidate molecules for inhibitors of the binding of TGF-beta binding-protein to a TGF-beta family member.

### 5 a. Combinatorial Peptide Libraries

Peptide molecules which are putative inhibitors of the binding of TGF-beta binding-protein to a TGF-beta family member may be obtained through the screening of combinatorial peptide libraries. Such libraries may either be prepared by one of skill in the art (*see e.g.*, U.S. Patent Nos. 4,528,266 and 4,359,535, and Patent Cooperation Treaty Publication Nos. WO 92/15679, WO 92/15677, WO 90/07862, WO 90/02809, or purchased from commercially available sources (*e.g.*, New England Biolabs Ph.D.<sup>TM</sup> Phage Display Peptide Library Kit).

### b. Antibodies

Antibodies which inhibit the binding of TGF-beta binding-protein to a TGF-beta family member may readily be prepared given the disclosure provided herein. Within the context of the present invention, antibodies are understood to include monoclonal antibodies, polyclonal antibodies, anti-idiotypic antibodies, antibody fragments (*e.g.*, Fab, and F(ab')<sub>2</sub>, F<sub>V</sub> variable regions, or complementarity determining regions). As discussed above, antibodies are understood to be specific against TGF-beta binding-protein, or against a specific TGF-beta family member, if they bind with a K<sub>a</sub> of greater than or equal to 10<sup>7</sup>M, preferably greater than or equal to 10<sup>8</sup>M, and do not bind to other TGF-beta binding-proteins, or, bind with a K<sub>a</sub> of less than or equal to 10<sup>6</sup>M. Furthermore, antibodies of the present invention should block or inhibit the binding of TGF-beta binding-protein to a TGF-beta family member.

The affinity of a monoclonal antibody or binding partner, as well as inhibition of binding can be readily determined by one of ordinary skill in the art (*see* Scatchard, *Ann. N.Y. Acad. Sci.* 51:660-672, 1949).

Briefly, polyclonal antibodies may be readily generated by one of ordinary skill in the art from a variety of warm-blooded animals such as horses, cows, various fowl, rabbits, mice, or rats. Typically, the TGF-beta binding-protein or unique peptide thereof of 13-20 amino acids (preferably conjugated to keyhole limpet hemocyanin by cross-linking with glutaraldehyde) is utilized to immunize the animal through intraperitoneal, intramuscular, intraocular, or subcutaneous injections, along with an adjuvant such as Freund's complete or incomplete adjuvant. Following several

booster immunizations, samples of serum are collected and tested for reactivity to the protein or peptide. Particularly preferred polyclonal antisera will give a signal on one of these assays that is at least three times greater than background. Once the titer of the animal has reached a plateau in terms of its reactivity to the protein, larger quantities of  
5 antisera may be readily obtained either by weekly bleedings, or by exsanguinating the animal.

Monoclonal antibodies may also be readily generated using conventional techniques (see U.S. Patent Nos. RE 32,011, 4,902,614, 4,543,439, and 4,411,993 which are incorporated herein by reference; see also *Monoclonal Antibodies, Hybridomas: A New Dimension in Biological Analyses*, Plenum Press, Kennett, McKearn, and Bechtol (eds.), 1980, and *Antibodies: A Laboratory Manual*, Harlow and Lane (eds.), Cold Spring Harbor Laboratory Press, 1988, which are also incorporated  
10 herein by reference).

Briefly, within one embodiment a subject animal such as a rat or mouse  
15 is immunized with TGF-beta binding-protein or portion thereof as described above. The protein may be admixed with an adjuvant such as Freund's complete or incomplete adjuvant in order to increase the resultant immune response. Between one and three weeks after the initial immunization the animal may be reimmunized with another booster immunization, and tested for reactivity to the protein utilizing assays described  
20 above. Once the animal has reached a plateau in its reactivity to the injected protein, it is sacrificed, and organs which contain large numbers of B cells such as the spleen and lymph nodes are harvested.

Cells which are obtained from the immunized animal may be immortalized by infection with a virus such as the Epstein-Barr virus (EBV) (see  
25 Glasky and Reading, *Hybridoma* 8(4):377-389, 1989). Alternatively, within a preferred embodiment, the harvested spleen and/or lymph node cell suspensions are fused with a suitable myeloma cell in order to create a "hybridoma" which secretes monoclonal antibody. Suitable myeloma lines include, for example, NS-1 (ATCC No. TIB 18), and P3X63 - Ag 8.653 (ATCC No. CRL 1580).

Following the fusion, the cells may be placed into culture plates  
30 containing a suitable medium, such as RPMI 1640, or DMEM (Dulbecco's Modified Eagles Medium) (JRH Biosciences, Lenexa, Kansas), as well as additional ingredients, such as fetal bovine serum (FBS, i.e., from Hyclone, Logan, Utah, or JRH Biosciences). Additionally, the medium should contain a reagent which selectively  
35 allows for the growth of fused spleen and myeloma cells such as HAT (hypoxanthine, aminopterin, and thymidine) (Sigma Chemical Co., St. Louis, Missouri). After about

seven days, the resulting fused cells or hybridomas may be screened in order to determine the presence of antibodies which are reactive against TGF-beta binding-protein (depending on the antigen used), and which block or inhibit the binding of TGF-beta binding-protein to a TGF-beta family member.

5 A wide variety of assays may be utilized to determine the presence of antibodies which are reactive against the proteins of the present invention, including for example countercurrent immuno-electrophoresis, radioimmunoassays, radioimmunoprecipitations, enzyme-linked immuno-sorbent assays (ELISA), dot blot assays, western blots, immunoprecipitation, inhibition or competition assays, and  
10 sandwich assays (*see* U.S. Patent Nos. 4,376,110 and 4,486,530; *see also* *Antibodies: A Laboratory Manual*, Harlow and Lane (eds.), Cold Spring Harbor Laboratory Press, 1988). Following several clonal dilutions and reassays, a hybridoma producing antibodies reactive against the desired protein may be isolated.

Other techniques may also be utilized to construct monoclonal  
15 antibodies (*see* William D. Huse et al., "Generation of a Large Combinational Library of the Immunoglobulin Repertoire in Phage Lambda," *Science* 246:1275-1281, December 1989; *see also* L. Sastry et al., "Cloning of the Immunological Repertoire in *Escherichia coli* for Generation of Monoclonal Catalytic Antibodies: Construction of a Heavy Chain Variable Region-Specific cDNA Library," *Proc. Natl. Acad. Sci. USA*  
20 86:5728-5732, August 1989; *see also* Michelle Alting-Mees et al., "Monoclonal Antibody Expression Libraries: A Rapid Alternative to Hybridomas," *Strategies in Molecular Biology* 3:1-9, January 1990). These references describe a commercial system available from Stratagene (La Jolla, California) which enables the production of antibodies through recombinant techniques. Briefly, mRNA is isolated from a B cell  
25 population, and utilized to create heavy and light chain immunoglobulin cDNA expression libraries in the  $\lambda$ ImmunoZap(H) and  $\lambda$ ImmunoZap(L) vectors. These vectors may be screened individually or co-expressed to form Fab fragments or antibodies (*see* Huse et al., *supra*; *see also* Sastry et al., *supra*). Positive plaques may subsequently be converted to a non-lytic plasmid which allows high level expression of  
30 monoclonal antibody fragments from *E. coli*.

Similarly, portions or fragments, such as Fab and Fv fragments, of antibodies may also be constructed utilizing conventional enzymatic digestion or recombinant DNA techniques to incorporate the variable regions of a gene which encodes a specifically binding antibody. Within one embodiment, the genes which  
35 encode the variable region from a hybridoma producing a monoclonal antibody of interest are amplified using nucleotide primers for the variable region. These primers

may be synthesized by one of ordinary skill in the art, or may be purchased from commercially available sources. Stratagene (La Jolla, California) sells primers for mouse and human variable regions including, among others, primers for  $V_{H\alpha}$ ,  $V_{H\beta}$ ,  $V_{H\gamma}$ ,  $V_{H\delta}$ ,  $C_{H1}$ ,  $V_L$  and  $C_L$  regions. These primers may be utilized to amplify heavy or light chain variable regions, which may then be inserted into vectors such as ImmunoZAP™ H or ImmunoZAP™ L (Stratagene), respectively. These vectors may then be introduced into *E. coli*, yeast, or mammalian-based systems for expression. Utilizing these techniques, large amounts of a single-chain protein containing a fusion of the  $V_H$  and  $V_L$  domains may be produced (*see* Bird et al., *Science* 242:423-426, 1988). In addition, such techniques may be utilized to change a "murine" antibody to a "human" antibody, without altering the binding specificity of the antibody.

Once suitable antibodies have been obtained, they may be isolated or purified by many techniques well known to those of ordinary skill in the art (*see Antibodies: A Laboratory Manual*, Harlow and Lane (eds.), Cold Spring Harbor Laboratory Press, 1988). Suitable techniques include peptide or protein affinity columns, HPLC or RP-HPLC, purification on protein A or protein G columns, or any combination of these techniques.

#### c. Mutant TGF-beta binding-proteins

As described herein and below in the Examples (e.g., Examples 8 and 9), altered versions of TGF-beta binding-protein which compete with native TGF-beta binding-protein's ability to block the activity of a particular TGF-beta family member should lead to increased bone density. Thus, mutants of TGF-beta binding-protein which bind to the TGF-beta family member but do not inhibit the function of the TGF-beta family member would meet the criteria. The mutant versions must effectively compete with the endogenous inhibitory functions of TGF-beta binding-protein.

#### d. Production of proteins

Although various genes (or portions thereof) have been provided herein, it should be understood that within the context of the present invention, reference to one or more of these genes includes derivatives of the genes that are substantially similar to the genes (and, where appropriate, the proteins (including peptides and polypeptides) that are encoded by the genes and their derivatives). As used herein, a nucleotide sequence is deemed to be "substantially similar" if: (a) the nucleotide sequence is derived from the coding region of the above-described genes and includes, for example, portions of the sequence or allelic variations of the sequences discussed



above, or alternatively, encodes a molecule which inhibits the binding of TGF-beta binding-protein to a member of the TGF-beta family, (b) the nucleotide sequence is capable of hybridization to nucleotide sequences of the present invention under moderate, high or very high stringency (see Sambrook et al., *Molecular Cloning: A Laboratory Manual*, 2nd ed., Cold Spring Harbor Laboratory Press, NY, 1989); or  
5 (c) the DNA sequences are degenerate as a result of the genetic code to the DNA sequences defined in (a) or (b). Further, the nucleic acid molecule disclosed herein includes both complementary and non-complementary sequences, provided the sequences otherwise meet the criteria set forth herein. Within the context of the present  
10 invention, high stringency means standard hybridization conditions (e.g., 5XSSPE, 0.5% SDS at 65°C, or the equivalent).

The structure of the proteins encoded by the nucleic acid molecules described herein may be predicted from the primary translation products using the hydrophobicity plot function of, for example, P/C Gene or Intelligenetics Suite  
15 (Intelligenetics, Mountain View, California), or according to the methods described by Kyte and Doolittle (*J. Mol. Biol.* 157:105-132, 1982).

Proteins of the present invention may be prepared in the form of acidic or basic salts, or in neutral form. In addition, individual amino acid residues may be modified by oxidation or reduction. Furthermore, various substitutions, deletions, or  
20 additions may be made to the amino acid or nucleic acid sequences, the net effect of which is to retain or further enhance or decrease the biological activity of the mutant or wild-type protein. Moreover, due to degeneracy in the genetic code, for example, there may be considerable variation in nucleotide sequences encoding the same amino acid sequence.

Other derivatives of the proteins disclosed herein include conjugates of  
25 the proteins along with other proteins or polypeptides. This may be accomplished, for example, by the synthesis of N-terminal or C-terminal fusion proteins which may be added to facilitate purification or identification of proteins (see U.S. Patent No. 4,851,341, see also, Hopp et al., *Bio/Technology* 6:1204, 1988.) Alternatively, fusion  
30 proteins such as Flag/TGF-beta binding-protein be constructed in order to assist in the identification, expression, and analysis of the protein.

Proteins of the present invention may be constructed using a wide variety of techniques described herein. Further, mutations may be introduced at particular loci by synthesizing oligonucleotides containing a mutant sequence, flanked  
35 by restriction sites enabling ligation to fragments of the native sequence. Following

ligation, the resulting reconstructed sequence encodes a derivative having the desired amino acid insertion, substitution, or deletion.

Alternatively, oligonucleotide-directed site-specific (or segment specific) mutagenesis procedures may be employed to provide an altered gene having particular codons altered according to the substitution, deletion, or insertion required. Exemplary methods of making the alterations set forth above are disclosed by Walder et al. (*Gene* 42:133, 1986); Bauer et al. (*Gene* 37:73, 1985); Craik (*BioTechniques*, January 1985, 12-19); Smith et al. (*Genetic Engineering: Principles and Methods*, Plenum Press, 1981); and Sambrook et al. (*supra*). Deletion or truncation derivatives of proteins (e.g., a soluble extracellular portion) may also be constructed by utilizing convenient restriction endonuclease sites adjacent to the desired deletion. Subsequent to restriction, overhangs may be filled in, and the DNA religated. Exemplary methods of making the alterations set forth above are disclosed by Sambrook et al. (*Molecular Cloning: A Laboratory Manual*, 2d Ed., Cold Spring Harbor Laboratory Press, 1989).

Mutations which are made in the nucleic acid molecules of the present invention preferably preserve the reading frame of the coding sequences. Furthermore, the mutations will preferably not create complementary regions that could hybridize to produce secondary mRNA structures, such as loops or hairpins, that would adversely affect translation of the mRNA. Although a mutation site may be predetermined, it is not necessary that the nature of the mutation *per se* be predetermined. For example, in order to select for optimum characteristics of mutants at a given site, random mutagenesis may be conducted at the target codon and the expressed mutants screened for indicative biological activity. Alternatively, mutations may be introduced at particular loci by synthesizing oligonucleotides containing a mutant sequence, flanked by restriction sites enabling ligation to fragments of the native sequence. Following ligation, the resulting reconstructed sequence encodes a derivative having the desired amino acid insertion, substitution, or deletion.

Nucleic acid molecules which encode proteins of the present invention may also be constructed utilizing techniques of PCR mutagenesis, chemical mutagenesis (Drinkwater and Klinedinst, *PNAS* 83:3402-3406, 1986), by forced nucleotide misincorporation (e.g., Liao and Wise *Gene* 88:107-111, 1990), or by use of randomly mutagenized oligonucleotides (Horwitz et al., *Genome* 3:112-117, 1989).

The present invention also provides for the manipulation and expression of the above described genes by culturing host cells containing a vector capable of expressing the above-described genes. Such vectors or vector constructs include either synthetic or cDNA-derived nucleic acid molecules encoding the desired protein, which

are operably linked to suitable transcriptional or translational regulatory elements. Suitable regulatory elements may be derived from a variety of sources, including bacterial, fungal, viral, mammalian, insect, or plant genes. Selection of appropriate regulatory elements is dependent on the host cell chosen, and may be readily  
5 accomplished by one of ordinary skill in the art. Examples of regulatory elements include: a transcriptional promoter and enhancer or RNA polymerase binding sequence, a transcriptional terminator, and a ribosomal binding sequence, including a translation initiation signal.

Nucleic acid molecules that encode any of the proteins described above  
10 may be readily expressed by a wide variety of prokaryotic and eukaryotic host cells, including bacterial, mammalian, yeast or other fungi, viral, insect, or plant cells. Methods for transforming or transfecting such cells to express foreign DNA are well known in the art (see, e.g., Itakura et al., U.S. Patent No. 4,704,362; Hinnen et al.,  
15 *Proc. Natl. Acad. Sci. USA* 75:1929-1933, 1978; Murray et al., U.S. Patent No. 4,801,542; Upshall et al., U.S. Patent No. 4,935,349; Hagen et al., U.S. Patent No. 4,784,950; Axel et al., U.S. Patent No. 4,399,216; Goeddel et al., U.S. Patent No. 4,766,075; and Sambrook et al. *Molecular Cloning: A Laboratory Manual*, 2nd ed., Cold Spring Harbor Laboratory Press, 1989; for plant cells see Czako and Marton,  
20 *Plant Physiol.* 104:1067-1071, 1994; and Paszkowski et al., *Biotech.* 24:387-392, 1992).

Bacterial host cells suitable for carrying out the present invention include *E. coli*, *B. subtilis*, *Salmonella typhimurium*, and various species within the genera *Pseudomonas*, *Streptomyces*, and *Staphylococcus*, as well as many other bacterial species well known to one of ordinary skill in the art. Representative  
25 examples of bacterial host cells include DH5 $\alpha$  (Stratagene, LaJolla, California).

Bacterial expression vectors preferably comprise a promoter which functions in the host cell, one or more selectable phenotypic markers, and a bacterial origin of replication. Representative promoters include the  $\beta$ -lactamase (penicillinase) and lactose promoter system (see Chang et al., *Nature* 275:615, 1978), the T7 RNA  
30 polymerase promoter (Studier et al., *Meth. Enzymol.* 185:60-89, 1990), the lambda promoter (Elvin et al., *Gene* 87:123-126, 1990), the *trp* promoter (Nichols and Yanofsky, *Meth. in Enzymology* 101:155, 1983) and the *tac* promoter (Russell et al., *Gene* 20:231, 1982). Representative selectable markers include various antibiotic resistance markers such as the kanamycin or ampicillin resistance genes. Many  
35 plasmids suitable for transforming host cells are well known in the art, including among others, pBR322 (see Bolivar et al., *Gene* 2:95, 1977), the pUC plasmids pUC18,

pUC19, pUC118, pUC119 (see Messing, *Meth. in Enzymology* 101:20-77, 1983 and Vieira and Messing, *Gene* 19:259-268, 1982), and pNH8A, pNH16a, pNH18a, and Bluescript M13 (Stratagene, La Jolla, California).

Yeast and fungi host cells suitable for carrying out the present invention include, among others, *Saccharomyces pombe*, *Saccharomyces cerevisiae*, the genera *Pichia* or *Kluyveromyces* and various species of the genus *Aspergillus* (McKnight et al., U.S. Patent No. 4,935,349). Suitable expression vectors for yeast and fungi include, among others, YCp50 (ATCC No. 37419) for yeast, and the amdS cloning vector pV3 (Turnbull, *Bio/Technology* 7:169, 1989), YRp7 (Struhl et al., *Proc. Natl. Acad. Sci. USA* 76:1035-1039, 1978), YEp13 (Broach et al., *Gene* 8:121-133, 1979), pJDB249 and pJDB219 (Beggs, *Nature* 275:104-108, 1978) and derivatives thereof.

Preferred promoters for use in yeast include promoters from yeast glycolytic genes (Hitzeman et al., *J. Biol. Chem.* 255:12073-12080, 1980; Alber and Kawasaki, *J. Mol. Appl. Genet.* 1:419-434, 1982) or alcohol dehydrogenase genes (Young et al., in *Genetic Engineering of Microorganisms for Chemicals*, Hollaender et al. (eds.), p. 355, Plenum, New York, 1982; Ammerer, *Meth. Enzymol.* 101:192-201, 1983). Examples of useful promoters for fungi vectors include those derived from *Aspergillus nidulans* glycolytic genes, such as the *adh3* promoter (McKnight et al., *EMBO J.* 4:2093-2099, 1985). The expression units may also include a transcriptional terminator. An example of a suitable terminator is the *adh3* terminator (McKnight et al., *ibid.*, 1985).

As with bacterial vectors, the yeast vectors will generally include a selectable marker, which may be one of any number of genes that exhibit a dominant phenotype for which a phenotypic assay exists to enable transformants to be selected. Preferred selectable markers are those that complement host cell auxotrophy, provide antibiotic resistance or enable a cell to utilize specific carbon sources, and include *leu2* (Broach et al., *ibid.*), *ura3* (Botstein et al., *Gene* 8:17, 1979), or *his3* (Struhl et al., *ibid.*). Another suitable selectable marker is the *cat* gene, which confers chloramphenicol resistance on yeast cells.

Techniques for transforming fungi are well known in the literature, and have been described, for instance, by Beggs (*ibid.*), Hinnen et al. (*Proc. Natl. Acad. Sci. USA* 75:1929-1933, 1978), Yelton et al. (*Proc. Natl. Acad. Sci. USA* 81:1740-1747, 1984), and Russell (*Nature* 301:167-169, 1983). The genotype of the host cell may contain a genetic defect that is complemented by the selectable marker present on the expression vector. Choice of a particular host and selectable marker is well within the level of ordinary skill in the art.

Protocols for the transformation of yeast are also well known to those of ordinary skill in the art. For example, transformation may be readily accomplished either by preparation of spheroplasts of yeast with DNA (*see* Hinnen et al., *PNAS USA* 75:1929, 1978) or by treatment with alkaline salts such as LiCl (*see* Itoh et al., *J. Bacteriology* 153:163, 1983). Transformation of fungi may also be carried out using polyethylene glycol as described by Cullen et al. (*BioTechnology* 5:369, 1987).

Viral vectors include those which comprise a promoter that directs the expression of an isolated nucleic acid molecule that encodes a desired protein as described above. A wide variety of promoters may be utilized within the context of the present invention, including for example, promoters such as MoMLV LTR, RSV LTR, Friend MuLV LTR, adenoviral promoter (Ohno et al., *Science* 265:781-784, 1994), neomycin phosphotransferase promoter/enhancer, late parvovirus promoter (Koering et al., *Hum. Gene Therap.* 5:457-463, 1994), Herpes TK promoter, SV40 promoter, metallothionein IIa gene enhancer/promoter, cytomegalovirus immediate early promoter, and the cytomegalovirus immediate late promoter. Within particularly preferred embodiments of the invention, the promoter is a tissue-specific promoter (*see e.g.*, WO 91/02805; EP 0,415,731; and WO 90/07936). Representative examples of suitable tissue specific promoters include neural specific enolase promoter, platelet derived growth factor beta promoter, bone morphogenic protein promoter, human alpha1-chimaerin promoter, synapsin I promoter and synapsin II promoter. In addition to the above-noted promoters, other viral-specific promoters (*e.g.*, retroviral promoters (including those noted above, as well as others such as HIV promoters), hepatitis, herpes (*e.g.*, EBV), and bacterial, fungal or parasitic (*e.g.*, malarial) -specific promoters may be utilized in order to target a specific cell or tissue which is infected with a virus, bacteria, fungus or parasite.

Mammalian cells suitable for carrying out the present invention include, among others COS, CHO, SaOS, osteosarcomas, KS483, MG-63, primary osteoblasts, and human or mammalian bone marrow stroma. Mammalian expression vectors for use in carrying out the present invention will include a promoter capable of directing the transcription of a cloned gene or cDNA. Preferred promoters include viral promoters and cellular promoters. Bone specific promoters include the bone sialo-protein and the promoter for osteocalcin. Viral promoters include the cytomegalovirus immediate early promoter (Boshart et al., *Cell* 41:521-530, 1985), cytomegalovirus immediate late promoter, SV40 promoter (Subramani et al., *Mol. Cell. Biol.* 1:854-864, 1981), MMTV LTR, RSV LTR, metallothionein-I, adenovirus E1a. Cellular promoters include the mouse metallothionein-I promoter (Palmiter et al., U.S. Patent No. 4,579,821), a

mouse V<sub>K</sub> promoter (Bergman et al., *Proc. Natl. Acad. Sci. USA* 81:7041-7045, 1983; Grant et al., *Nucl. Acids Res.* 15:5496, 1987) and a mouse V<sub>H</sub> promoter (Loh et al., *Cell* 33:85-93, 1983). The choice of promoter will depend, at least in part, upon the level of expression desired or the recipient cell line to be transfected.

5           Such expression vectors may also contain a set of RNA splice sites located downstream from the promoter and upstream from the DNA sequence encoding the peptide or protein of interest. Preferred RNA splice sites may be obtained from adenovirus and/or immunoglobulin genes. Also contained in the expression vectors is a polyadenylation signal located downstream of the coding sequence of interest. Suitable  
10 polyadenylation signals include the early or late polyadenylation signals from SV40 (Kaufman and Sharp, *ibid.*), the polyadenylation signal from the Adenovirus 5 E1B region and the human growth hormone gene terminator (DeNoto et al., *Nuc. Acids Res.* 9:3719-3730, 1981). The expression vectors may include a noncoding viral leader sequence, such as the Adenovirus 2 tripartite leader, located between the promoter and  
15 the RNA splice sites. Preferred vectors may also include enhancer sequences, such as the SV40 enhancer. Expression vectors may also include sequences encoding the adenovirus VA RNAs. Suitable expression vectors can be obtained from commercial sources (e.g., Stratagene, La Jolla, California).

          Vector constructs comprising cloned DNA sequences can be introduced  
20 into cultured mammalian cells by, for example, calcium phosphate-mediated transfection (Wigler et al., *Cell* 14:725, 1978; Corsaro and Pearson, *Somatic Cell Genetics* 7:603, 1981; Graham and Van der Eb, *Virology* 52:456, 1973), electroporation (Neumann et al., *EMBO J.* 1:841-845, 1982), or DEAE-dextran mediated transfection (Ausubel et al. (eds.), *Current Protocols in Molecular Biology*, John Wiley and Sons,  
25 Inc., NY, 1987). To identify cells that have stably integrated the cloned DNA, a selectable marker is generally introduced into the cells along with the gene or cDNA of interest. Preferred selectable markers for use in cultured mammalian cells include genes that confer resistance to drugs, such as neomycin, hygromycin, and methotrexate. The selectable marker may be an amplifiable selectable marker. Preferred amplifiable  
30 selectable markers are the DHFR gene and the neomycin resistance gene. Selectable markers are reviewed by Thilly (*Mammalian Cell Technology*, Butterworth Publishers, Stoneham, Massachusetts, which is incorporated herein by reference).

          Mammalian cells containing a suitable vector are allowed to grow for a period of time, typically 1-2 days, to begin expressing the DNA sequence(s) of interest.  
35 Drug selection is then applied to select for growth of cells that are expressing the selectable marker in a stable fashion. For cells that have been transfected with an

amplifiable, selectable marker the drug concentration may be increased in a stepwise manner to select for increased copy number of the cloned sequences, thereby increasing expression levels. Cells expressing the introduced sequences are selected and screened for production of the protein of interest in the desired form or at the desired level.

- 5 Cells that satisfy these criteria can then be cloned and scaled up for production.

Protocols for the transfection of mammalian cells are well known to those of ordinary skill in the art. Representative methods include calcium phosphate mediated transfection, electroporation, lipofection, retroviral, adenoviral and protoplast fusion-mediated transfection (*see* Sambrook et al., *supra*). Naked vector constructs can  
10 also be taken up by muscular cells or other suitable cells subsequent to injection into the muscle of a mammal (or other animals).

Numerous insect host cells known in the art can also be useful within the present invention, in light of the subject specification. For example, the use of baculoviruses as vectors for expressing heterologous DNA sequences in insect cells has  
15 been reviewed by Atkinson et al. (*Pestic. Sci.* 28:215-224, 1990).

Numerous plant host cells known in the art can also be useful within the present invention, in light of the subject specification. For example, the use of *Agrobacterium rhizogenes* as vectors for expressing genes in plant cells has been reviewed by Sinkar et al. (*J. Biosci. (Bangalore)* 11:47-58, 1987).

20 Within related aspects of the present invention, proteins of the present invention may be expressed in a transgenic animal whose germ cells and somatic cells contain a gene which encodes the desired protein and which is operably linked to a promoter effective for the expression of the gene. Alternatively, in a similar manner transgenic animals may be prepared that lack the desired gene (*e.g.*, "knock-out" mice).  
25 Such transgenics may be prepared in a variety of non-human animals, including mice, rats, rabbits, sheep, dogs, goats and pigs (*see* Hammer et al., *Nature* 315:680-683, 1985, Palmiter et al., *Science* 222:809-814, 1983, Brinster et al., *Proc. Natl. Acad. Sci. USA* 82:4438-4442, 1985, Palmiter and Brinster, *Cell* 41:343-345, 1985, and U.S. Patent Nos. 5,175,383, 5,087,571, 4,736,866, 5,387,742, 5,347,075, 5,221,778, and  
30 5,175,384). Briefly, an expression vector, including a nucleic acid molecule to be expressed together with appropriately positioned expression control sequences, is introduced into pronuclei of fertilized eggs, for example, by microinjection. Integration of the injected DNA is detected by blot analysis of DNA from tissue samples. It is preferred that the introduced DNA be incorporated into the germ line of the animal so  
35 that it is passed on to the animal's progeny. Tissue-specific expression may be achieved through the use of a tissue-specific promoter, or through the use of an

inducible promoter, such as the metallothionein gene promoter (Palmiter et al., 1983, *ibid*), which allows regulated expression of the transgene.

Proteins can be isolated by, among other methods, culturing suitable host and vector systems to produce the recombinant translation products of the present invention. Supernatants from such cell lines, or protein inclusions or whole cells where the protein is not excreted into the supernatant, can then be treated by a variety of purification procedures in order to isolate the desired proteins. For example, the supernatant may be first concentrated using commercially available protein concentration filters, such as an Amicon or Millipore Pellicon ultrafiltration unit. Following concentration, the concentrate may be applied to a suitable purification matrix such as, for example, an anti-protein antibody bound to a suitable support. Alternatively, anion or cation exchange resins may be employed in order to purify the protein. As a further alternative, one or more reverse-phase high performance liquid chromatography (RP-HPLC) steps may be employed to further purify the protein. Other methods of isolating the proteins of the present invention are well known in the skill of the art.

A protein is deemed to be "isolated" within the context of the present invention if no other (undesired) protein is detected pursuant to SDS-PAGE analysis followed by Coomassie blue staining. Within other embodiments, the desired protein can be isolated such that no other (undesired) protein is detected pursuant to SDS-PAGE analysis followed by silver staining.

### 3. Nucleic Acid Molecules

Within other aspects of the invention, nucleic acid molecules are provided which are capable of inhibiting TGF-beta binding-protein binding to a member of the TGF-beta family. For example, within one embodiment antisense oligonucleotide molecules are provided which specifically inhibit expression of TGF-beta binding-protein nucleic acid sequences (*see generally*, Hirashima et al. in *Molecular Biology of RNA: New Perspectives* (M. Inouye and B. S. Dudock, eds., 1987 Academic Press, San Diego, p. 401); *Oligonucleotides: Antisense Inhibitors of Gene Expression* (J.S. Cohen, ed., 1989 MacMillan Press, London); Stein and Cheng, *Science* 261:1004-1012, 1993; WO 95/10607; U.S. Patent No. 5,359,051; WO 92/06693; and EP-A2-612844). Briefly, such molecules are constructed such that they are complementary to, and able to form Watson-Crick base pairs with, a region of transcribed TGF-beta binding-protein mRNA sequence. The resultant double-stranded



nucleic acid interferes with subsequent processing of the mRNA, thereby preventing protein synthesis (see Example 10).

Within other aspects of the invention, ribozymes are provided which are capable of inhibiting the TGF-beta binding-protein binding to a member of the TGF-beta family. As used herein, "ribozymes" are intended to include RNA molecules that contain anti-sense sequences for specific recognition, and an RNA-cleaving enzymatic activity. The catalytic strand cleaves a specific site in a target RNA at greater than stoichiometric concentration. A wide variety of ribozymes may be utilized within the context of the present invention, including for example, the hammerhead ribozyme (for example, as described by Forster and Symons, *Cell* 48:211-220, 1987; Haseloff and Gerlach, *Nature* 328:596-600, 1988; Walbot and Bruening, *Nature* 334:196, 1988; Haseloff and Gerlach, *Nature* 334:585, 1988); the hairpin ribozyme (for example, as described by Haseloff et al., U.S. Patent No. 5,254,678, issued October 19, 1993 and Hempel et al., European Patent Publication No. 0 360 257, published March 26, 1990); and *Tetrahymena* ribosomal RNA-based ribozymes (see Cech et al., U.S. Patent No. 4,987,071). Ribozymes of the present invention typically consist of RNA, but may also be composed of DNA, nucleic acid analogs (e.g., phosphorothioates), or chimerics thereof (e.g., DNA/RNA/RNA).

#### 4. Labels

The gene product or any of the candidate molecules described above and below, may be labeled with a variety of compounds, including for example, fluorescent molecules, toxins, and radionuclides. Representative examples of fluorescent molecules include fluorescein, *Phycobili* proteins, such as phycoerythrin, rhodamine, Texas red and luciferase. Representative examples of toxins include ricin, abrin diphtheria toxin, cholera toxin, gelonin, pokeweed antiviral protein, tritin, *Shigella* toxin, and *Pseudomonas* exotoxin A. Representative examples of radionuclides include Cu-64, Ga-67, Ga-68, Zr-89, Ru-97, Tc-99m, Rh-105, Pd-109, In-111, I-123, I-125, I-131, Re-186, Re-188, Au-198, Au-199, Pb-203, At-211, Pb-212 and Bi-212. In addition, the antibodies described above may also be labeled or conjugated to one partner of a ligand binding pair. Representative examples include avidin-biotin, and riboflavin-riboflavin binding protein.

Methods for conjugating or labeling the molecules described herein with the representative labels set forth above may be readily accomplished by one of ordinary skill in the art (see Trichothecene Antibody Conjugate, U.S. Patent No. 4,744,981; Antibody Conjugate, U.S. Patent No. 5,106,951; Fluorogenic Materials and

- Labeling Techniques, U.S. Patent No. 4,018,884; Metal Radionuclide Labeled Proteins for Diagnosis and Therapy, U.S. Patent No. 4,897,255; and Metal Radionuclide Chelating Compounds for Improved Chelation Kinetics, U.S. Patent No. 4,988,496; *see also* Inman, *Methods In Enzymology*, Vol. 34, *Affinity Techniques, Enzyme Purification: Part B*, Jakoby and Wilchek (eds.), Academic Press, New York, p. 30, 1974; *see also* Wilchek and Bayer, "The Avidin-Biotin Complex in Bioanalytical Applications," *Anal. Biochem.* 171:1-32, 1988).

#### PHARMACEUTICAL COMPOSITIONS

- As noted above, the present invention also provides a variety of pharmaceutical compositions, comprising one of the above-described molecules which inhibits the TGF-beta binding-protein binding to a member of the TGF-beta family along with a pharmaceutically or physiologically acceptable carrier, excipients or diluents. Generally, such carriers should be nontoxic to recipients at the dosages and concentrations employed. Ordinarily, the preparation of such compositions entails combining the therapeutic agent with buffers, antioxidants such as ascorbic acid, low molecular weight (less than about 10 residues) polypeptides, proteins, amino acids, carbohydrates including glucose, sucrose or dextrans, chelating agents such as EDTA, glutathione and other stabilizers and excipients. Neutral buffered saline or saline mixed with nonspecific serum albumin are exemplary appropriate diluents.

- In addition, the pharmaceutical compositions of the present invention may be prepared for administration by a variety of different routes. In addition, pharmaceutical compositions of the present invention may be placed within containers, along with packaging material which provides instructions regarding the use of such pharmaceutical compositions. Generally, such instructions will include a tangible expression describing the reagent concentration, as well as within certain embodiments, relative amounts of excipient ingredients or diluents (*e.g.*, water, saline or PBS) which may be necessary to reconstitute the pharmaceutical composition.

#### METHODS OF TREATMENT

- The present invention also provides methods for increasing the mineral content and mineral density of bone. Briefly, numerous conditions result in the loss of bone mineral content, including for example, disease, genetic predisposition, accidents which result in the lack of use of bone (*e.g.*, due to fracture), therapeutics which effect bone resorption, or which kill bone forming cells and normal aging. Through use of the molecules described herein which inhibit the TGF-beta binding-protein binding to a

TGF-beta family member such conditions may be treated or prevented. As utilized herein, it should be understood that bone mineral content has been increased, if bone mineral content has been increased in a statistically significant manner (*e.g.*, greater than one-half standard deviation), at a selected site.

5 A wide variety of conditions which result in loss of bone mineral content may be treated with the molecules described herein. Patients with such conditions may be identified through clinical diagnosis utilizing well known techniques (see, *e.g.*, Harrison's Principles of Internal Medicine, McGraw-Hill, Inc.). Representative examples of diseases that may be treated included dysplasias, wherein there is abnormal  
10 growth or development of bone. Representative examples of such conditions include achondroplasia, cleidocranial dysostosis, enchondromatosis, fibrous dysplasia, Gaucher's, hypophosphatemic rickets, Marfan's, multiple hereditary exotoses, neurofibromatosis, osteogenesis imperfecta, osteopetrosis, osteopoikilosis, sclerotic lesions, fractures, periodontal disease, pseudoarthrosis and pyogenic osteomyelitis.

15 Other conditions which may be treated or prevented include a wide variety of causes of osteopenia (*i.e.*, a condition that causes greater than one standard deviation of bone mineral content or density below peak skeletal mineral content at youth). Representative examples of such conditions include anemic states, conditions caused steroids, conditions caused by heparin, bone marrow disorders, scurvy,  
20 malnutrition, calcium deficiency, idiopathic osteoporosis, congenital osteopenia or osteoporosis, alcoholism, chronic liver disease, senility, postmenopausal state, oligomenorrhea, amenorrhea, pregnancy, diabetes mellitus, hyperthyroidism, Cushing's disease, acromegaly, hypogonadism, immobilization or disuse, reflex sympathetic dystrophy syndrome, transient regional osteoporosis and osteomalacia.

25 Within one aspect of the present invention, bone mineral content or density may be increased by administering to a warm-blooded animal a therapeutically effective amount of a molecule which inhibits the TGF-beta binding-protein binding to a TGF-beta family member. Examples of warm-blooded animals that may be treated include both vertebrates and mammals, including for example horses, cows, pigs,  
30 sheep, dogs, cats, rats and mice. Representative examples of therapeutic molecules include ribozymes, ribozyme genes, antisense oligonucleotides and antibodies (*e.g.*, humanized antibodies).

Within other aspects of the present invention, methods are provided for increasing bone density, comprising the step of introducing into cells which home to  
35 bone a vector which directs the expression of a molecule which inhibits the TGF-beta binding-protein binding to a member of the TGF-beta family, and administering the

vector containing cells to a warm-blooded animal. Briefly, cells which home to bone may be obtained directly from the bone of patients (*e.g.*, cells obtained from the bone marrow such as CD34+, osteoblasts, osteocytes, and the like), from peripheral blood, or from cultures.

5 A vector which directs the expression of a molecule that inhibits the TGF-beta binding-protein binding to a member of the TGF-beta family is introduced into the cells. Representative examples of suitable vectors include viral vectors such as herpes viral vectors (*e.g.*, U.S. Patent No. 5,288,641), adenoviral vectors (*e.g.*, WO 94/26914, WO 93/9191; Kolls et al., *PNAS* 91(1):215-219, 1994; Kass-Eisler et al.,  
10 *PNAS* 90(24):11498-502, 1993; Guzman et al., *Circulation* 88(6):2838-48, 1993; Guzman et al., *Cir. Res.* 73(6):1202-1207, 1993; Zabner et al., *Cell* 75(2):207-216, 1993; Li et al., *Hum Gene Ther.* 4(4):403-409, 1993; Caillaud et al., *Eur. J. Neurosci.* 5(10):1287-1291, 1993; Vincent et al., *Nat. Genet.* 5(2):130-134, 1993; Jaffe et al., *Nat. Genet.* 1(5):372-378, 1992; and Levrero et al., *Gene* 101(2):195-202, 1991), adeno-  
15 associated viral vectors (WO 95/13365; Flotte et al., *PNAS* 90(22):10613-10617, 1993), baculovirus vectors, parvovirus vectors (Koering et al., *Hum. Gene Therap.* 5:457-463, 1994), pox virus vectors (Panicali and Paoletti, *PNAS* 79:4927-4931, 1982; and Ozaki et al., *Biochem. Biophys. Res. Comm.* 193(2):653-660, 1993), and retroviruses (*e.g.*, EP 0,415,731; WO 90/07936; WO 91/0285, WO 94/03622; WO 93/25698; WO 93/25234; U.S. Patent No. 5,219,740; WO 93/11230; WO 93/10218). Viral  
20 vectors may likewise be constructed which contain a mixture of different elements (*e.g.*, promoters, envelope sequences and the like) from different viruses, or non-viral sources. Within various embodiments, either the viral vector itself, or a viral particle which contains the viral vector may be utilized in the methods and compositions  
25 described below.

Within other embodiments of the invention, nucleic acid molecules which encode a molecule which inhibits the TGF-beta binding-protein binding to a member of the TGF-beta family themselves may be administered by a variety of techniques, including, for example, administration of asialosomucoid (ASOR) conjugated with poly-L-lysine DNA complexes (Cristano et al., *PNAS* 92:122-92126,  
30 1993), DNA linked to killed adenovirus (Curiel et al., *Hum. Gene Ther.* 3(2):147-154, 1992), cytofectin-mediated introduction (DMRIE-DOPE, Vical, California), direct DNA injection (Acsadi et al., *Nature* 352:815-818, 1991); DNA ligand (Wu et al., *J. of Biol. Chem.* 264:16985-16987, 1989); lipofection (Felgner et al., *Proc. Natl. Acad. Sci. USA* 84:7413-7417, 1989); liposomes (Pickering et al., *Circ.* 89(1):13-21, 1994; and  
35 Wang et al., *PNAS* 84:7851-7855, 1987); microprojectile bombardment (Williams

et al., *PNAS* 88:2726-2730, 1991); and direct delivery of nucleic acids which encode the protein itself either alone (Vile and Hart, *Cancer Res.* 53: 3860-3864, 1993), or utilizing PEG-nucleic acid complexes.

Representative examples of molecules which may be expressed by the  
5 vectors of present invention include ribozymes and antisense molecules, each of which are discussed in more detail above.

Determination of increased bone mineral content may be determined directly through the use of X-rays (e.g., Dual Energy X-ray Absorptometry or "DEXA"), or by inference through bone turnover markers (osteoblast specific alkaline  
10 phosphatase, osteocalcin, type 1 procollagen C' propeptide (PICP), and total alkaline phosphatase; see Comier, C., *Curr. Opin. in Rheu.* 7:243, 1995), or markers of bone resorption (pyridinoline, deoxypyridinoline, N-telopeptide, urinary hydroxyproline, plasma tartrate-resistant acid phosphatases and galactosyl hydroxylysine; see Comier,  
15 *supra*). The amount of bone mass may also be calculated from body weights, or utilizing other methods (see Guinness-Hey, *Metab. Bone Dis. and Rel. Res.* 5:177-181, 1984).

As will be evident to one of skill in the art, the amount and frequency of administration will depend, of course, on such factors as the nature and severity of the indication being treated, the desired response, the condition of the patient, and so forth.  
20 Typically, the compositions may be administered by a variety of techniques, as noted above.

The following examples are offered by way of illustration, and not by way of limitation.

## EXAMPLES

EXAMPLE 1

## SCLEROSTEOSIS MAPS TO THE LONG ARM OF HUMAN CHROMOSOME 17

Genetic mapping of the defect responsible for sclerosteosis in humans  
5 localized the gene responsible for this disorder to the region of human chromosome 17  
that encodes a novel TGF-beta binding-protein family member. In sclerosteosis,  
skeletal bone displays a substantial increase in mineral density relative to that of  
unaffected individuals. Bone in the head displays overgrowth as well. Sclerosteosis  
patients are generally healthy although they may exhibit variable degrees of syndactyly  
10 at birth and variable degrees of cranial compression and nerve compression in the skull.

Linkage analysis of the gene defect associated with sclerosteosis was  
conducted by applying the homozygosity mapping method to DNA samples collected  
from 24 South African Afrikaaner families in which the disease occurred. (Sheffield  
et al., 1994, *Human Molecular Genetics* 3:1331-1335. "Identification of a Bardet-Biedl  
15 syndrome locus on chromosome 3 and evaluation of an efficient approach to  
homozygosity mapping"). The Afrikaaner population of South Africa is genetically  
homogeneous; the population is descended from a small number of founders who  
colonized the area several centuries ago, and it has been isolated by geographic and  
social barriers since the founding. Sclerosteosis is rare everywhere in the world outside  
20 the Afrikaaner community, which suggests that a mutation in the gene was present in  
the founding population and has since increased in numbers along with the increase in  
the population. The use of homozygosity mapping is based on the assumption that  
DNA mapping markers adjacent to a recessive mutation are likely to be homozygous  
in affected individuals from consanguineous families and isolated populations.

25 A set of 371 microsatellite markers (Research Genetics, Set 6) from the  
autosomal chromosomes was selected to type pools of DNA from sclerosteosis patient  
samples. The DNA samples for this analysis came from 29 sclerosteosis patients in 24  
families, 59 unaffected family members and a set of unrelated control individuals from  
the same population. The pools consisted of 4-6 individuals, either affected  
30 individuals, affected individuals from consanguineous families, parents and unaffected  
siblings, or unrelated controls. In the pools of unrelated individuals and in most of the  
pools with affected individuals or family members analysis of the markers showed  
several allele sizes for each marker. One marker, D17S1299, showed an indication of  
homozygosity: one band in several of the pools of affected individuals.

All 24 sclerosteosis families were typed with a total of 19 markers in the region of D17S1299 (at 17q12-q21). Affected individuals from every family were shown to be homozygous in this region, and 25 of the 29 individuals were homozygous for a core haplotype; they each had the same alleles between D17S1787 and D17S930. The other four individuals had one chromosome which matched this haplotype and a second which did not. In sum, the data compellingly suggested that this 3 megabase region contained the sclerosteosis mutation. Sequence analysis of most of the exons in this 3 megabase region identified a nonsense mutation in the novel TGF-beta binding-protein coding sequence (C>T mutation at position 117 of Sequence ID No. 1 results in a stop codon). This mutation was shown to be unique to sclerosteosis patients and carriers of Afrikaaner descent. The identity of the gene was further confirmed by identifying a mutation in its intron (A>T mutation at position +3 of the intron) which results in improper mRNA processing in a single, unrelated patient with diagnosed sclerosteosis.

## EXAMPLE 2

### TISSUE-SPECIFICITY OF TGF-BETA BINDING-PROTEIN GENE EXPRESSION

#### A. Human Beer Gene Expression by RT-PCR:

First-strand cDNA was prepared from the following total RNA samples using a commercially available kit ("Superscript Preamplification System for First-Strand cDNA Synthesis", Life Technologies, Rockville, MD): human brain, human liver, human spleen, human thymus, human placenta, human skeletal muscle, human thyroid, human pituitary, human osteoblast (NHOst from Clonetics Corp., San Diego, CA), human osteosarcoma cell line (Saos-2, ATCC# HTB-85), human bone, human bone marrow, human cartilage, vervet monkey bone, saccharomyces cerevisiae, and human peripheral blood monocytes. All RNA samples were purchased from a commercial source (Clontech, Palo Alto, CA), except the following which were prepared in-house: human osteoblast, human osteosarcoma cell line, human bone, human cartilage and vervet monkey bone. These in-house RNA samples were prepared using a commercially available kit ("TRI Reagent", Molecular Research Center, Inc., Cincinnati, OH).

PCR was performed on these samples, and additionally on a human genomic sample as a control. The sense Beer oligonucleotide primer had the sequence 5'-CCGGAGCTGGAGAACAACAAG-3' (SEQ ID NO:19). The antisense Beer oligonucleotide primer had the sequence 5'-GCACTGGCCGGAGCACACC-3' (SEQ

ID NO:20). In addition, PCR was performed using primers for the human beta-actin gene, as a control. The sense beta-actin oligonucleotide primer had the sequence 5'-AGGCCAACCGCGAGAAGATGA CC -3' (SEQ ID NO:21). The antisense beta-actin oligonucleotide primer had the sequence 5'-GAAGT CCAGGGCGACGTAGCA-3' (SEQ ID NO:22). PCR was performed using standard conditions in 25 ul reactions, with an annealing temperature of 61 degrees Celsius. Thirty-two cycles of PCR were performed with the Beer primers and twenty-four cycles were performed with the beta-actin primers.

Following amplification, 12 ul from each reaction were analyzed by agarose gel electrophoresis and ethidium bromide staining. See Figure 2A.

#### B. RNA In-situ Hybridization of Mouse Embryo Sections:

The full length mouse *Beer* cDNA (Sequence ID No. 11) was cloned into the pCR2.1 vector (Invitrogen, Carlsbad, CA) in the antisense and sense direction using the manufacturer's protocol. <sup>35</sup>S-alpha-GTP-labeled cRNA sense and antisense transcripts were synthesized using in-vitro transcription reagents supplied by Ambion, Inc (Austin, TX). In-situ hybridization was performed according to the protocols of Lyons et al. (*J. Cell Biol.* 111:2427-2436, 1990).

The mouse *Beer* cRNA probe detected a specific message expressed in the neural tube, limb buds, blood vessels and ossifying cartilages of developing mouse embryos. Panel A in Figure 3 shows expression in the apical ectodermal ridge (aer) of the limb (l) bud, blood vessels (bv) and the neural tube (nt). Panel B shows expression in the 4<sup>th</sup> ventricle of the brain (4). Panel C shows expression in the mandible (ma) cervical vertebrae (cv), occipital bone (oc), palate (pa) and a blood vessel (bv). Panel D shows expression in the ribs (r) and a heart valve (va). Panel A is a transverse section of 10.5 dpc embryo. Panel B is a sagittal section of 12.5 dpc embryo and panels C and D are sagittal sections of 15.5 dpc embryos.

ba=branchial arch, h=heart, te=telencephalon (forebrain), b=brain, f=frontonasal mass, g=gut, h=heart, j=jaw, li=liver, lu=lung, ot=otic vesicle, ao=, sc=spinal cord, skm=skeletal muscle, ns=nasal sinus, th=thymus, to=tongue, fl=forelimb, di=diaphragm

### EXAMPLE 3

#### 35 EXPRESSION AND PURIFICATION OF RECOMBINANT BEER PROTEIN

##### A. Expression in COS-1 Cells:



The DNA sequence encoding the full length human Beer protein was amplified using the following PCR oligonucleotide primers: The 5' oligonucleotide primer had the sequence 5'-**AAGCTTGGTACCATGCAGCTCCCAC**-3' (SEQ ID NO:23) and contained a HindIII restriction enzyme site (in bold) followed by 19 nucleotides of the *Beer* gene starting 6 base pairs prior to the presumed amino terminal start codon (ATG). The 3' oligonucleotide primer had the sequence 5'-**AAGCTTCTACTTGTCATCGTCGTCCT** TGTAGTCGTAGGCGTTCTCCAGCT-3' (SEQ ID NO:24) and contained a HindIII restriction enzyme site (in bold) followed by a reverse complement stop codon (CTA) followed by the reverse complement of the FLAG epitope (underlined, Sigma-Aldrich Co., St. Louis, MO) flanked by the reverse complement of nucleotides coding for the carboxy terminal 5 amino acids of the Beer. The PCR product was TA cloned ("Original TA Cloning Kit", Invitrogen, Carlsbad, CA) and individual clones were screened by DNA sequencing. A sequence-verified clone was then digested by HindIII and purified on a 1.5% agarose gel using a commercially available reagents ("QIAquick Gel Extraction Kit", Qiagen Inc., Valencia, CA). This fragment was then ligated to HindIII digested, phosphatase-treated pcDNA3.1 (Invitrogen, Carlsbad, CA) plasmid with T4 DNA ligase. DH10B *E. coli* were transformed and plated on LB, 100 µg/ml ampicillin plates. Colonies bearing the desired recombinant in the proper orientation were identified by a PCR-based screen, using a 5' primer corresponding to the T7 promoter/priming site in pcDNA3.1 and a 3' primer with the sequence 5'- GCACTGGCCGGAGCACACC-3' (SEQ ID NO:25) that corresponds to the reverse complement of internal BEER sequence. The sequence of the cloned fragment was confirmed by DNA sequencing.

COS-1 cells (ATCC# CRL-1650) were used for transfection. 50 µg of the expression plasmid pcDNA-Beer-Flag was transfected using a commercially available kit following protocols supplied by the manufacturer ("DEAE-Dextran Transfection Kit", Sigma Chemical Co., St. Louis, MO). The final media following transfection was DMEM (Life Technologies, Rockville, MD) containing 0.1% Fetal Bovine Serum. After 4 days in culture, the media was removed. Expression of recombinant BEER was analyzed by SDS-PAGE and Western Blot using anti-FLAG M2 monoclonal antibody (Sigma-Aldrich Co., St. Louis, MO). Purification of recombinant BEER protein was performed using an anti-FLAG M2 affinity column ("Mammalian Transient Expression System", Sigma-Aldrich Co., St. Louis, MO). The column profile was analyzed via SDS-PAGE and Western Blot using anti-FLAG M2 monoclonal antibody.

B. Expression in SF9 insect cells:

The human *Beer* gene sequence was amplified using PCR with standard conditions and the following primers:

Sense primer: 5'-GTCGTCGGATCCATGGGGTGGCAGGCGTTCAAGAATGAT-3'  
5 (SEQ ID NO:26)

Antisense primer: 5'-GTCGTCAAGCTTCTACTTGTCATCGTCCTTGTAGTCGTA  
GGCGTTCTCCAGCTCGGC-3' (SEQ ID NO:27)

The resulting cDNA contained the coding region of Beer with two modifications. The N-terminal secretion signal was removed and a FLAG epitope tag  
10 (Sigma) was fused in frame to the C-terminal end of the insert. BamHI and HindIII cloning sites were added and the gene was subcloned into pMelBac vector (Invitrogen) for transfer into a baculoviral expression vector using standard methods.

Recombinant baculoviruses expressing Beer protein were made using the Bac-N-Blue transfection kit (Invitrogen) and purified according to the manufacturers  
15 instructions.

SF9 cells (Invitrogen) were maintained in TNM\_FH media (Invitrogen) containing 10% fetal calf serum. For protein expression, SF9 cultures in spinner flasks were infected at an MOI of greater than 10. Samples of the media and cells were taken daily for five days, and Beer expression monitored by western blot using an anti-FLAG  
20 M2 monoclonal antibody (Sigma) or an anti-Beer rabbit polyclonal antiserum.

After five days the baculovirus-infected SF9 cells were harvested by centrifugation and cell associated protein was extracted from the cell pellet using a high salt extraction buffer (1.5 M NaCl, 50 mM Tris pH 7.5). The extract (20 ml per 300 ml culture) was clarified by centrifugation, dialyzed three times against four liters of  
25 Tris buffered saline (150 mM NaCl, 50 mM Tris pH 7.5), and clarified by centrifugation again. This high salt fraction was applied to Hitrap Heparin (Pharmacia; 5 ml bed volume), washed extensively with HEPES buffered saline (25 mM HEPES 7.5, 150 mM NaCl) and bound proteins were eluted with a gradient from 150 mM NaCl to 1200 mM NaCl. Beer elution was observed at approximately 800 mM NaCl. Beer  
30 containing fractions were supplemented to 10% glycerol and 1 mM DTT and frozen at -80 degrees C.

#### EXAMPLE 4

35 PREPARATION AND TESTING OF POLYCLONAL ANTIBODIES TO BEER, GREMLIN, AND  
DAN

A. Preparation of antigen:

The DNA sequences of Human *Beer*, Human *Gremlin*, and Human *Dan* were amplified using standard PCR methods with the following oligonucleotide primers:

5 H. Beer

Sense: 5' -GACTTGGATCCCAGGGGTGGCAGGCGTTC- 3' (SEQ ID NO:28)

Antisense 5' -AGCATAAGCTTCTAGTAGGCGTTCTCCAG- 3' (SEQ ID NO:29)

H. Gremlin

Sense: 5' -GACTTGGATCCGAAGGGAAAAAGAAAGGG- 3' (SEQ ID NO:30)

10 Antisense: 5' -AGCATAAGCTTTTAATCCAAATCGATGGA- 3' (SEQ ID NO:31)

H. Dan

Sense: 5' -ACTACGAGCTCGGCCCCACCACCCATCAACAAG- 3' (SEQ ID NO:32)

Antisense: 5' -ACTTAGAAGCTTTCAGTCCTCAGCCCCCTCTTCC-3' (SEQ ID NO:33)

15 In each case the listed primers amplified the entire coding region minus the secretion signal sequence. These include restriction sites for subcloning into the bacterial expression vector pQE-30 (Qiagen Inc., Valencia, CA) at sites BamHI/HindIII for Beer and Gremlin, and sites SacI/HindIII for Dan. pQE30 contains a coding sequence for a 6x His tag at the 5' end of the cloning region. The completed constructs  
20 were transformed into *E. coli* strain M-15/pRep (Qiagen Inc) and individual clones verified by sequencing. Protein expression in M-15/pRep and purification (6xHis affinity tag binding to Ni-NTA coupled to Sepharose) were performed as described by the manufacturer (Qiagen, The QIAexpressionist).

25 The *E. coli*-derived Beer protein was recovered in significant quantity using solubilization in 6M guanidine and dialyzed to 2-4M to prevent precipitation during storage. Gremlin and Dan protein were recovered in higher quantity with solubilization in 6M guanidine and a post purification guanidine concentration of 0.5M.

B. Production and testing of polyclonal antibodies:

30 Polyclonal antibodies to each of the three antigens were produced in rabbit and in chicken hosts using standard protocols (R & R Antibody, Stanwood, WA; standard protocol for rabbit immunization and antisera recovery; Short Protocols in Molecular Biology, 2nd edition, 1992, 11.37- 11.41. Contributors Helen M. Cooper and Yvonne Paterson; chicken antisera was generated with Strategic Biosolutions,  
35 Ramona, CA).

Rabbit antisera and chicken egg IgY fraction were screened for activity

via Western blot. Each of the three antigens was separated by PAGE and transferred to 0.45um nitrocellulose (Novex, San Diego, CA). The membrane was cut into strips with each strip containing approximately 75 ng of antigen. The strips were blocked in 3% Blotting Grade Block (Bio-Rad Laboratories, Hercules, CA) and washed 3 times in 1X Tris buffer saline (TBS) /0.02% TWEEN buffer. The primary antibody (preimmunization bleeds, rabbit antisera or chicken egg IgY in dilutions ranging from 1:100 to 1:10,000 in blocking buffer) was incubated with the strips for one hour with gentle rocking. A second series of three washes 1X TBS/0.02%TWEEN was followed by an one hour incubation with the secondary antibody (peroxidase conjugated donkey anti-rabbit, Amersham Life Science, Piscataway, NJ; or peroxidase conjugated donkey anti-chicken, Jackson ImmunoResearch, West Grove, PA). A final cycle of 3X washes of 1X TBS/0.02%TWEEN was performed and the strips were developed with Lumi-Light Western Blotting Substrate (Roche Molecular Biochemicals, Mannheim, Germany).

15

C. Antibody cross-reactivity test:

Following the protocol described in the previous section, nitrocellulose strips of Beer, Gremlin or Dan were incubated with dilutions (1:5000 and 1:10,000) of their respective rabbit antisera or chicken egg IgY as well as to antisera or chicken egg IgY (dilutions 1:1000 and 1:5000) made to the remaining two antigens. The increased levels of nonmatching antibodies was performed to detect low affinity binding by those antibodies that may be seen only at increased concentration. The protocol and duration of development is the same for all three binding events using the protocol described above. There was no antigen cross-reactivity observed for any of the antigens tested.

25

### EXAMPLE 5

#### INTERACTION OF BEER WITH TGF-BETA SUPER-FAMILY PROTEINS

The interaction of Beer with proteins from different phylogenetic arms of the TGF- $\beta$  superfamily were studied using immunoprecipitation methods. Purified TGF $\beta$ -1, TGF $\beta$ -2, TGF $\beta$ -3, BMP-4, BMP-5, BMP-6 and GDNF were obtained from commercial sources (R&D systems; Minneapolis, MN). A representative protocol is as follows. Partially purified Beer was dialyzed into HEPES buffered saline (25 mM HEPES 7.5, 150 mM NaCl). Immunoprecipitations were done in 300 ul of IP buffer (150 mM NaCl, 25 mM Tris pH 7.5, 1mM EDTA, 1.4 mM  $\beta$ -mercaptoethanol, 0.5 % triton X 100, and 10% glycerol). 30 ng recombinant human BMP-5 protein (R&D

30

35

systems) was applied to 15 ul of FLAG affinity matrix (Sigma; St Louis MO) in the presence and absence of 500 ng FLAG epitope-tagged Beer. The proteins were incubated for 4 hours @ 4°C and then the affinity matrix-associated proteins were washed 5 times in IP buffer (1 ml per wash). The bound proteins were eluted from the affinity matrix in 60 microliters of 1X SDS PAGE sample buffer. The proteins were resolved by SDS PAGE and Beer associated BMP-5 was detected by western blot using anti-BMP-5 antiserum (Research Diagnostics, Inc) (see Figure 5).

#### BEER Ligand Binding Assay:

FLAG-Beer protein (20 ng) is added to 100 ul PBS/0.2% BSA and adsorbed into each well of 96 well microtiter plate previously coated with anti-FLAG monoclonal antibody (Sigma; St Louis MO) and blocked with 10% BSA in PBS. This is conducted at room temperature for 60 minutes. This protein solution is removed and the wells are washed to remove unbound protein. BMP-5 is added to each well in concentrations ranging from 10 pM to 500 nM in PBS/0.2% BSA and incubated for 2 hours at room temperature. The binding solution is removed and the plate washed with three times with 200ul volumes of PBS/0.2% BSA. BMP-5 levels are then detected using BMP-5 anti-serum via ELISA (F.M. Ausubel et al (1998) Current Protocols in Mol Biol. Vol 2 11.2.1-11.2.22). Specific binding is calculated by subtracting non-specific binding from total binding and analyzed by the LIGAND program (Munson and Podbard, Anal. Biochem., 107, p220-239, (1980).

In a variation of this method, Beer is engineered and expressed as a human Fc fusion protein. Likewise the ligand BMP is engineered and expressed as mouse Fc fusion. These proteins are incubated together and the assay conducted as described by Mellor et al using homogeneous time resolved fluorescence detection (G.W. Mellor et al., *J of Biomol Screening*, 3(2) 91-99, 1998).

#### EXAMPLE 6

##### SCREENING ASSAY FOR INHIBITION OF TGF-BETA BINDING-PROTEIN BINDING TO TGF-BETA FAMILY MEMBERS

The assay described above is replicated with two exceptions. First, BMP concentration is held fixed at the K<sub>d</sub> determined previously. Second, a collection of antagonist candidates is added at a fixed concentration (20 uM in the case of the small organic molecule collections and 1 uM in antibody studies). These candidate molecules (antagonists) of TGF-beta binding-protein binding include organic

compounds derived from commercial or internal collections representing diverse chemical structures. These compounds are prepared as stock solutions in DMSO and are added to assay wells at  $\leq 1\%$  of final volume under the standard assay conditions. These are incubated for 2 hours at room temperature with the BMP and Beer, the solution removed and the bound BMP is quantitated as described. Agents that inhibit 40% of the BMP binding observed in the absence of compound or antibody are considered antagonists of this interaction. These are further evaluated as potential inhibitors based on titration studies to determine their inhibition constants and their influence on TGF-beta binding-protein binding affinity. Comparable specificity control assays may also be conducted to establish the selectivity profile for the identified antagonist through studies using assays dependent on the BMP ligand action (e.g. BMP/BMP receptor competition study).

#### EXAMPLE 7

##### INHIBITION OF TGF-BETA BINDING-PROTEIN LOCALIZATION TO BONE MATRIX

Evaluation of inhibition of localization to bone matrix (hydroxyapatite) is conducted using modifications to the method of Nicolas (Nicolas, V. *Calcif Tissue Int* 57:206, 1995). Briefly,  $^{125}\text{I}$ -labelled TGF-beta binding-protein is prepared as described by Nicolas (*supra*). Hydroxyapatite is added to each well of a 96 well microtiter plate equipped with a polypropylene filtration membrane (Polyfiltroninc, Weymouth MA). TGF-beta binding-protein is added to 0.2% albumin in PBS buffer. The wells containing matrix are washed 3 times with this buffer. Adsorbed TGF-beta binding-protein is eluted using 0.3M NaOH and quantitated.

Inhibitor identification is conducted via incubation of TGF-beta binding-protein with test molecules and applying the mixture to the matrix as described above. The matrix is washed 3 times with 0.2% albumin in PBS buffer. Adsorbed TGF-beta binding-protein is eluted using 0.3 M NaOH and quantitated. Agents that inhibit 40% of the TGF-beta binding-protein binding observed in the absence of compound or antibody are considered bone localization inhibitors. These inhibitors are further characterized through dose response studies to determine their inhibition constants and their influence on TGF-beta binding-protein binding affinity.

### EXAMPLE 8

#### CONSTRUCTION OF TGF-BETA BINDING-PROTEIN MUTANT

##### A. Mutagenesis:

5 A full-length TGF-beta binding-protein cDNA in pBluescript SK serves as a template for mutagenesis. Briefly, appropriate primers (see the discussion provided above) are utilized to generate the DNA fragment by polymerase chain reaction using Vent DNA polymerase (New England Biolabs, Beverly, MA). The polymerase chain reaction is run for 23 cycles in buffers provided by the manufacturer using a 57°C annealing temperature. The product is then exposed to two restriction enzymes and after isolation using agarose gel electrophoresis, ligated back into pRBP4-503 from which the matching sequence has been removed by enzymatic digestion. Integrity of the mutant is verified by DNA sequencing.

##### B. Mammalian Cell Expression and Isolation of Mutant TGF-beta binding-protein:

15 The mutant TGF-beta binding-protein cDNAs are transferred into the pcDNA3.1 mammalian expression vector described in EXAMPLE 3. After verifying the sequence, the resultant constructs are transfected into COS-1 cells, and secreted protein is purified as described in EXAMPLE 3.

20

### EXAMPLE 9

#### ANIMAL MODELS -I

##### GENERATION OF TRANSGENIC MICE OVEREXPRESSING THE *BEER* GENE

25 The ~200 kilobase (kb) BAC clone 15G5, isolated from the CITB mouse genomic DNA library (distributed by Research Genetics, Huntsville, AL) was used to determine the complete sequence of the mouse *Beer* gene and its 5' and 3' flanking regions. A 41 kb SalI fragment, containing the entire gene body, plus ~17 kb of 5' flanking and ~20 kb of 3' flanking sequence was sub-cloned into the BamHI site of the SuperCosI cosmid vector (Stratagene, La Jolla, CA) and propagated in the *E. coli* strain 30 DH10B. From this cosmid construct, a 35 kb MluI - AvIII restriction fragment (Sequence No. 6), including the entire mouse *Beer* gene, as well as 17 kb and 14 kb of 5' and 3' flanking sequence, respectively, was then gel purified, using conventional means, and used for microinjection of mouse zygotes (DNX Transgenics; US Patent No. 4,873,191). Founder animals in which the cloned DNA fragment was integrated 35 randomly into the genome were obtained at a frequency of 5-30% of live-born pups. The presence of the transgene was ascertained by performing Southern blot analysis of

genomic DNA extracted from a small amount of mouse tissue, such as the tip of a tail DNA was extracted using the following protocol: tissue was digested overnight at 55° C in a lysis buffer containing 200 mM NaCl, 100 mM Tris pH8.5, 5 mM EDTA, 0.2% SDS and 0.5 mg/ml Proteinase K. The following day, the DNA was extracted once with phenol/chloroform (50:50), once with chloroform/isoamylalcohol (24:1) and precipitated with ethanol. Upon resuspension in TE (10mM Tris pH7.5, 1 mM EDTA) 8-10 ug of each DNA sample were digested with a restriction endonuclease, such as EcoRI, subjected to gel electrophoresis and transferred to a charged nylon membrane, such as HyBondN+ (Amersham, Arlington Heights, IL ). The resulting filter was then hybridized with a radioactively labelled fragment of DNA deriving from the mouse *Beer* gene locus, and able to recognize both a fragment from the endogenous gene locus and a fragment of a different size deriving from the transgene. Founder animals were bred to normal non-transgenic mice to generate sufficient numbers of transgenic and non-transgenic progeny in which to determine the effects of *Beer* gene overexpression. For these studies, animals at various ages (for example, 1 day, 3 weeks, 6 weeks, 4 months) are subjected to a number of different assays designed to ascertain gross skeletal formation, bone mineral density, bone mineral content, osteoclast and osteoblast activity, extent of endochondral ossification, cartilage formation, etc. The transcriptional activity from the transgene may be determined by extracting RNA from various tissues, and using an RT-PCR assay which takes advantage of single nucleotide polymorphisms between the mouse strain from which the transgene is derived (129Sv/J) and the strain of mice used for DNA microinjection [(C57BL5/J x SJL/J)F2].

#### ANIMAL MODELS - II

##### 25      DISRUPTION OF THE MOUSE BEER GENE BY HOMOLOGOUS RECOMBINATION

Homologous recombination in embryonic stem (ES) cells can be used to inactivate the endogenous mouse *Beer* gene and subsequently generate animals carrying the loss-of-function mutation. A reporter gene, such as the *E. coli*  $\beta$ -galactosidase gene, was engineered into the targeting vector so that its expression is controlled by the endogenous *Beer* gene's promoter and translational initiation signal. In this way, the spatial and temporal patterns of *Beer* gene expression can be determined in animals carrying a targeted allele.

The targeting vector was constructed by first cloning the drug-selectable phosphoglycerate kinase (PGK) promoter driven *neomycin-resistance* gene (*neo*) cassette from pGT-N29 (New England Biolabs, Beverly, MA) into the cloning vector pSP72 (Promega, Madison, WI). PCR was used to flank the PGK*neo* cassette with



bacteriophage P1 loxP sites, which are recognition sites for the P1 Cre recombinase (Hoess et al., PNAS USA, 79:3398, 1982). This allows subsequent removal of the neo-resistance marker in targeted ES cells or ES cell-derived animals (US Patent 4,959,317). The PCR primers were comprised of the 34 nucleotide (ntd) loxP  
 5 sequence, 15-25 ntd complementary to the 5' and 3' ends of the PGKneo cassette, as well as restriction enzyme recognition sites (BamHI in the sense primer and EcoRI in the anti-sense primer) for cloning into pSP72. The sequence of the sense primer was 5'-

AATCTGGATCCATAACTTCGTATAGCATACATTATACGAAGTTATCTGCAG  
 10 GATTCGAGGGCCCCCT-3' (SEQ ID NO:34); sequence of the anti-sense primer was 5'-AATCTGAATTCCACCGGTGTTAATTAAATAACTTCGT  
 ATAATGTATGCTATACGAAGTTATAGATCTAGAG TCAGCTTCTGA-3' (SEQ ID NO:35).

The next step was to clone a 3.6 kb XhoI-HindIII fragment, containing  
 15 the *E. coli*  $\beta$ -galactosidase gene and SV40 polyadenylation signal from pSV $\beta$  (Clontech, Palo Alto, CA) into the pSP72-PGKneo plasmid. The "short arm" of homology from the mouse *Beer* gene locus was generated by amplifying a 2.4 kb fragment from the BAC clone 15G5. The 3' end of the fragment coincided with the translational initiation site of the *Beer* gene, and the anti-sense primer used in the PCR  
 20 also included 30 ntd complementary to the 5' end of the  $\beta$ -galactosidase gene so that its coding region could be fused to the Beer initiation site in-frame. The approach taken for introducing the "short arm" into the pSP72- $\beta$ gal-PGKneo plasmid was to linearize the plasmid at a site upstream of the  $\beta$ -gal gene and then to co-transform this fragment with the "short arm" PCR product and to select for plasmids in which the PCR product  
 25 was integrated by homologous recombination. The sense primer for the "short arm" amplification included 30 ntd complementary to the pSP72 vector to allow for this recombination event. The sequence of the sense primer was 5'-ATTTAGGTGACACT ATAGAACTCGAGCAGCTGAAGCTTAACCACATGGTGGCTCACAACCAT-3' (SEQ ID NO:36) and the sequence of the anti-sense primer was 5'-  
 30 AACGACGGCCAGTGAATCCGTA ATCATGGTCATGCTGCCAGGTGGAGGAGGGCA-3' (SEQ ID NO:37).

The "long arm" from the *Beer* gene locus was generated by amplifying a 6.1 kb fragment from BAC clone 15G5 with primers which also introduce the rare-cutting restriction enzyme sites SgrAI, FseI, Ascl and PacI. Specifically, the sequence  
 35 of the sense primer was 5'-ATTACCACCGGTGACACCCGCTTCCTGACAG-3' (SEQ ID NO:38); the sequence of the anti-sense primer was 5'-

ATTACTTAATTAAACATGGCGCGCCAT  
 ATGGCCGGCCCCCTAATTGCGGCGCATCGTTAATT-3' (SEQ ID NO:39). The resulting PCR product was cloned into the TA vector (Invitrogen, Carlsbad, CA ) as an intermediate step.

5 The mouse *Beer* gene targeting construct also included a second selectable marker, the *herpes simplex virus 1 thymidine kinase* gene (HSVTK) under the control of rous sarcoma virus long terminal repeat element (RSV LTR). Expression of this gene renders mammalian cells sensitive (and inviable) to gancyclovir; it is therefore a convenient way to select against neomycin-resistant cells in which the  
 10 construct has integrated by a non-homologous event (US Patent 5,464,764). The RSVLTR-HSVTK cassette was amplified from pPS1337 using primers that allow subsequent cloning into the FseI and Ascl sites of the "long arm"-TA vector plasmid. For this PCR, the sequence of the sense primer was 5'-  
 15 ATTACGGCCGGCCGCAAAGGAATTCAAGA TCTGA-3' (SEQ ID NO:40); the sequence of the anti-sense primer was 5'-ATTACGGCGCGCCCCCTC  
 ACAGGCCGCACCCAGCT-3' (SEQ ID NO:41).

The final step in the construction of the targeting vector involved cloning the 8.8 kb SgrAI-Ascl fragment containing the "long arm" and RSVLTR-HSVTK gene into the SgrAI and Ascl sites of the pSP72-"short arm"- $\beta$ gal-PGKneo  
 20 plasmid. This targeting vector was linearized by digestion with either Ascl or PacI before electroporation into ES cells.

### EXAMPLE 10

#### 25 ANTISENSE-MEDIATED BEER INACTIVATION

17-nucleotide antisense oligonucleotides are prepared in an overlapping format, in such a way that the 5' end of the first oligonucleotide overlaps the translation  
 30 initiating AUG of the Beer transcript, and the 5' ends of successive oligonucleotides occur in 5 nucleotide increments moving in the 5' direction (up to 50 nucleotides away), relative to the Beer AUG. Corresponding control oligonucleotides are designed and prepared using equivalent base composition but redistributed in sequence to inhibit any significant hybridization to the coding mRNA. Reagent delivery to the test cellular  
 35 system is conducted through cationic lipid delivery (P.L. Felgner, *Proc. Natl. Acad. Sci. USA* 84:7413, 1987). 2 ug of antisense oligonucleotide is added to 100 ul of reduced serum media (Opti-MEM 1 reduced serum media; Life Technologies, Gaithersburg MD) and this is mixed with Lipofectin reagent (6 ul) (Life Technologies,

Gaithersburg MD) in the 100 ul of reduced serum media. These are mixed, allowed to complex for 30 minutes at room temperature and the mixture is added to previously seeded MC3T3E21 or KS483 cells. These cells are cultured and the mRNA recovered. Beer mRNA is monitored using RT-PCR in conjunction with Beer specific primers. In addition, separate experimental wells are collected and protein levels characterized through western blot methods described in Example 4. The cells are harvested, resuspended in lysis buffer (50 mM Tris pH 7.5, 20 mM NaCl, 1mM EDTA, 1% SDS) and the soluble protein collected. This material is applied to 10-20 % gradient denaturing SDS PAGE. The separated proteins are transferred to nitrocellulose and the western blot conducted as above using the antibody reagents described. In parallel, the control oligonucleotides are added to identical cultures and experimental operations are repeated. Decrease in Beer mRNA or protein levels are considered significant if the treatment with the antisense oligonucleotide results in a 50% change in either instance compared to the control scrambled oligonucleotide. This methodology enables selective gene inactivation and subsequent phenotype characterization of the mineralized nodules in the tissue culture model.





TTTCTTTCTCTCTCTAATTTTAAATTGTAAAGAAAAAATTTTAAACAGAAAGACATTAATATTAAGAGCTTCTGAGCAATT  
GCTCTTTTCTTTGGCAATTCTTCAAGCTGGGCACTTCTTCAAAAGATGAAATAGTCTGCTTTTAAAGAGTTAAAGTTACAT  
ATTTATTTTCTCACTTAAGTTATTTTACGCAAAATTTTCTTTCTTAAGAGATGAAGATGTTAATATTCTTTTATGAATTAC  
CACTCTTTCTTCAAGAGTCCAGAGACATTCTTAATTAAGAGCAATGAATCATGTAAGAGAG

5

Sequence ID No. 4: Truncated Human Beer protein from Sclerosteosis

MOLEFLEBYLLUNHTAFRIVE\*

10

Sequence ID No. 5: Human BEER cDNA encoding protein variant (V10I)

[illegible]







[illegible]

5

Sequence ID No. 11:      Mouse BEER cDNA    (coding region only)

[illegible]

20 Sequence ID No. 12: Mouse BEER protein (complete sequence)

MQFSLAFCLICLLVHAFPCWEGQGQWQAFRNDATEVIFSLGEYFEFFENNQTMINRADINGGREHHFYDANDVSEYSRE  
 LHYTRFLTDGFCRSAPFVTELVCSGQGQGFARLLEHIGAVFWWRNGEDFACIFDRYRAGQVQLLCFGAGAPRSRKVLV  
 ASQCHRLTRFHHQSELKDFGFETARFQHGKHPFGAGAKHHQSELENAV

25

Sequence ID No. 13: Rat BEER cDNA (complete coding region plus 5' UTR)

30 GAGGACCGAGTGCCTTCTCTCTTCTGGACCCATGACAGCTCTCACTAGCGCCCTTGCCCTGCGCTGCTTGTATACATGCA  
GCCTTCTTTGCTGTGGAGAGCCAGGGGTGGCAAGCTTCAAGAAATGATGCCACAGAAATCATCCCGGGAATCAGAGAGTA  
CCCGAGAGCTTCTCAGGAACCTAGAGAACACACAGACCATGAAACCGGGCCGAGAACGGAGGCGAGACCCGCCACCATCCTT  
ATGACACCAAGACGTGTCCGASTACAGCTGCCCGAGAGTCACTACACCGCTTCTGTACCGAGCGGCGCTGTCCGCAAT  
GCCAAGCGGTACCGAGTTGTTGTCTCGGGGCAATGGGGCCCGCGCGCTGCTGCCCAAGGCCATCGGGCGCGTGA  
35 GTGGTGGCGCCGAACGGACCCGACTTCCGCTGCATCCCGGATCGCTACCGCGCGCAGCGGGTGCAGCTCTGTGCCCG  
GCGCGCGCGCGCGCTCGCGCAAGGTGCGTCTGGTGGCTGTGCAAGTSCAAGCGCCTCACCGCTTCCACACCCAG









5  
10  
15  
20  
25  
30  
35

5  
10  
15  
20  
25  
30  
35

[illegible]





[illegible]

5  
10  
15  
20  
25  
30  
35



[illegible]

5  
 10  
 15  
 20  
 25  
 30  
 35

[illegible]

Sequence ID No. 18: Human Beer Genomic Sequence (This gene has two exons, at positions 161-427 abd 3186-5219).

20 tagaggagaa gtctttgggg agggtttgct ctgagcacac ccctttccct ccctccgggg 60  
ctgaggggaaa catgggacca gccctgcccc agcctgtcct cattggctgg catgaagcag 120  
25 agaggggctt taaaaaggcg accgtgtctc ggctggagac cagagcctgt gctactggaa 180  
ggtagggctgc cctcctctgg ctggtaccat gcagctccca ctggccctgt gtctcgtctg 240  
cctgctggta cacacagcct tccgtgtagt ggagggccag gggtaggcagg cgttcaagaa 300  
30 tgatgccacg gaaatcatcc ccgagctcgg agagtacccc gagcctccac cggagctgga 360  
gaacaacaag accatgaacc gggcggagaa cggagggcgg cctccccacc acccctttga 420  
35 gaccaaaggt atggggtgga ggagagaatt cttagtaaaa gatcctgggg aggttttaga 480

aacttctctt tgggaggctt ggaagactgg ggtagacca gtgaagattg ctggcctctg 540  
ccagcactgg tcgaggaaca gtcttgctg gaggtggggg aagaatggct cgctgggtgca 600  
5 gccttcaaat tcaggtgcag aggcattgagg caacagacgc tggtagagagc ccagggcagg 660  
gaggacgctg ggggtggtag ggtatggcat cagggcatca gaacaggctc aggggctcag 720  
aaaagaaaag gtttcaaaga atctcctctt gggaatatag gagccacgtc cagctgctgg 780  
10 taccactggg aagggaaaca ggtaaggag cctcccatcc acagaacagc acctgtgggg 840  
caccggacac tctatgctgg tggtagctgt cccaccaca cagaccaca tcatggaatc 900  
15 cccaggaggt gaacccccag ctgaagggg aagaaacagg ttccaggcac tcagtaactt 960  
ggtagtgaga agagctgagg tgtgaacctg gtttgatcca actgcaagat agccctgggtg 1020  
tgtggggggg tgtgggggac agatctccac aaagcagtgg ggaggaaggc cagagaggca 1080  
20 cccctgcagt gtgcattgcc catggcctgc ccaggagct ggcacttgaa ggaatgggag 1140  
ttttcggcac agtttttagcc cctgacatgg gtgcagctga gtccaggccc tggaggggag 1200  
25 agcagcatcc tctgtgcagg agtagggaca tctgtcctca gcagccacc cagtcccaac 1260  
cttgccatcat tccaggggag ggagaaggaa gaggaaccct gggttcctgg tcaggcctgc 1320  
acagagaagc ccaggtgaca gtgtgcatct ggctctataa ttggcaggaa tcctgaggcc 1380  
30 atgggggctg ctgaaatgac acttcagact aagagcttcc ctgtcctctg gccattatcc 1440  
aggtggcaga gaagtccact gccaggctc ctggaccca gccctccccg cctcacaacc 1500  
35 tgttgggact atggggtgct aaaaagggca actgcatggg aggccagcca ggaccctccg 1560



tcttcaaaat ggaggacaag ggggctccc cccacagctc ccttcttagg caaggtcagc 1620  
tgggctccag cgactgcctg aagggctgta aggaacccaa acacaaaatg tccaccttgc 1680  
5 tggactccca cgagaggcca cagccctga ggaagccaca tgctcaaac aaagtcatga 1740  
tctgcagagg aagtgcctgg cctaggggcg ctattctcga aaagccgcaa aatgccccct 1800  
tccctgggca aatgcccccc tgaccacaca cacattccag cctgcagag gtgaggatgc 1860  
10 aaaccagccc acagaccaga aagcagcccc agacgatggc agtggccaca tctccccctgc 1920  
tgtgcttgct cttcagagtg ggggtggggg gtggccttct ctgtccctc tctggtttgg 1980  
15 tcttaagact atttttcatt ctttcttgc acattggaac tatccccatg aaacctttgg 2040  
gggtggactg gtactcacac gacgaccagc tatttaaaaa gctccccacc atctaagtcc 2100  
accataggag acatggtaa ggtgtgtgca ggggatcagg ccaggcctcg gagcccaatc 2160  
20 tctgctgcc cagggagtat caccatgagg cgccattca gataacacag aacaagaaat 2220  
gtgccagca gagagccagg tcaatgtttg tggcagctga acctgtaggt tttgggtcag 2280  
25 agctcagggc cctatggta ggaaagtaac gacagtaaaa agcagccctc agtccatcc 2340  
cccagcccag cctcccatgg atgctgaac gcagagcctc cactcttgcc ggagccaaaa 2400  
gggtctggga ccccaggaa gtggagtccg gagatgcagc ccagcctttt gggcaagtcc 2460  
30 ttttctctgg ctgggcctca gtattctcat tgataatgag ggggttggac aactgcctt 2520  
tgattccttt caagtcta atgaattcctgt cctgatcacc tccccttcag tccctgcct 2580  
35 ccacagcagc tgccctgatt tattacctc aattaacctc tactcctttc tccatcccc 2640

gtccacccct cccaagtggc tggaaaagga atttgggaga agccagagcc aggcagaagg 2700  
tgtgctgagt acttaccctg cccaggccag ggaccctgcg gcacaagtgt ggcttaaata 2760  
5 ataagaagac cccagaagag aaatgataat aataatacat aacagccgac gctttcagct 2820  
atatgtgcc aatggatatt tctgcattgc gtgtgtaatg gattaactcg caatgcttgg 2880  
ggcggcccat tttgcagaca ggaagaagag agaggtaag gaacttgccc aagatgacac 2940  
10 ctgcagtgag cgatggagcc ctgggtgtttg aaccccagca gtcatttggc tccgagggga 3000  
caggggtgcgc aggagagctt tccaccagct ctagagcatc tgggaccttc ctgcaataga 3060  
15 tgttcagggg caaaagcctc tggagacagg ctgggcaaaa gcagggctgg ggtggagaga 3120  
gacgggcccgg tccagggcag ggggtggccag gcgggcccgc accctcacgc gcgcctctct 3180  
ccacagacgt gtccgagtac agctgccgcg agctgcactt caccgctac gtgaccgatg 3240  
20 ggccgtgccg cagcgccaag ccggtcaccg agctgggtgtg ctccggccag tgcggcccgg 3300  
cgcgcctgct gcccacgcc atcggccgcg gcaagtgggtg gcgacctagt gggcccgact 3360  
25 tccgctgcat ccccgaccgc taccgcgcg agcgcgtgca gctgctgtgt cccggtggtg 3420  
aggcgccgcg cgcgcgcaag gtgcgcctgg tggcctcgtg caagtgcaag cgcctcacc 3480  
gcttccacaa ccagtcggag ctcaaggact tcgggaccga ggccgctcgg ccgcagaagg 3540  
30 gccggaagcc gcggccccgc gcccgagcg ccaaagccaa ccaggccgag ctggagaacg 3600  
cctactagag cccgcccgcg cccctcccca ccggcgggcg ccccgccct gaaccgcgc 3660  
35 cccacatttc tgcctctgc gcgtgggttg attgtttata ttctattgta aatgcctgca 3720

acccagggca gggggctgag accttccagg ccttgaggaa tcccgggccc cggcaaggcc 3780  
cccttcagcc cgccagctga ggggtcccac ggggcagggg aggggaattga gagtcacaga 3840  
5 cactgagcca cgcagccccg cctctggggc cgcttacctt tgctgggtccc acttcagagg 3900  
aggcagaaat ggaagcattt tcaccgcctt ggggttttaa gggagcgggtg tgggagtggg 3960  
aaagtccagg gactgggttaa gaaagttaga taagattccc ccttgcacct cgctgcccatt 4020  
10 cagaaagcct gaggcgtgcc cagagcacia gactgggggc aactgtagat gtgggtttcta 4080  
gtcctggctc tgccactaac ttgctgtgta accttgaact acacaattct ccttcgggac 4140  
15 ctcaatttcc actttgtaaa atgaggggtgg aggtgggaat aggatctcga ggagactatt 4200  
ggcatatgat tccaaggact ccagtgcctt ttgaatgggc agaggtgaga gagagagaga 4260  
gaaagagaga gaatgaatgc agttgcattg attcagtgcc aaggtcactt ccagaattca 4320  
20 gagttgtgat gctctcttct gacagccaaa gatgaaaaac aaacagaaaa aaaaaagtaa 4380  
agagtctatt tatggctgac atatttacgg ctgacaaact cctggaagaa gctatgctgc 4440  
25 tccccagcct ggcttccccg gatgtttggc tacctccacc cctccatctc aaagaaataa 4500  
catcatccat tggggtagaa aaggagaggg tccgaggggtg gtgggagggg tagaaatcac 4560  
atccgccccca acttcccaaa gaggcagatc cctccccga cccatagcca tgttttaaag 4620  
30 tcaccttccg aagagaagtg aaaggttcaa ggacactggc cttgcaggcc cgagggagca 4680  
gccatcacia actcacagac cagcacatcc cttttgagac accgccttct gccaccact 4740  
35 cacggacaca tttctgccta gaaaacagct tcttactgct cttacatgtg atggcatatc 4800

ttacactaaa agaataattat tgggggaaaa actacaagtg ctgtacatat gctgagaaac 4860  
tgcagagcat aatagctgcc acccaaaaaat ctttttgaaa atcatttcca gacaacctct 4920  
5 tactttctgt gtagttttta attgttaaaa aaaaaaagtt ttaaacagaa gcacatgaca 4980  
tatgaaagcc tgcaggactg gtcgtttttt tggcaattct tccacgtggg acttgtccac 5040  
aagaatgaaa gtagtgggtt ttaaagagtt aagttacata tttattttct cacttaagtt 5100  
10 atttatgcaa aagtttttct tgtagagaat gacaatgtta atattgcttt atgaactaac 5160  
agtctgttct tccagagtcc agagacattg ttaataaaga caatgaatca tgaccgaaag 5220  
15 gatgtggtct cattttgtca accacacatg acgtcatttc tgtcaaagtt gacaccttc 5280  
tcttggtcac tagagctcca accttggaca cacctttgac tgctctctgg tggcccttgt 5340  
ggcaattatg tcttcctttg aaaagtcatg tttatccctt cctttccaaa cccagaccgc 5400  
20 atttcttcac ccagggcatg gtaataacct cagccttgta tccttttagc agcctccct 5460  
ccatgctggc ttccaaaatg ctgttctcat tgtatcactc cctgctcaa aagccttcca 5520  
25 tagctcccc ttgccagga tcaagtgcag tttccctatc tgacatggga ggccttctct 5580  
gcttgactcc cacctccac tccaccaagc ttcctactga ctccaaatgg tcatgcagat 5640  
cctgcttcc ttagtttgcc atccacactt agcaccccca ataactaatc ctctttcttt 5700  
30 aggattcaca ttacttgta tctcttcccc taaccttcca gagatgttcc aatctcccat 5760  
gatccctctc tcctctgagg ttccagcccc ttttgtctac accactactt tggttcctaa 5820  
35 ttctgttttc catttgacag tcattcatgg aggaccagcc tggccaagtc ctgcttagta 5880

ctggcataga caacacaaag ccaagtacaa ttcaggacca gctcacagga aacttcacct 5940  
tcttcgaagt gtggatttga tgcctcctgg gtagaaatgt aggatcttca aaagtgggcc 6000  
5 agcctcctgc acttctctca aagtctcgcc tccccaaagt gtcttaatag tgctggatgc 6060  
tagctgagtt agcatcttca gatgaagagt aacctaaag ttactcttca gtgcccctaa 6120  
ggtagggatgg tcaactggaa agctttaaat taagtccagc ctaccttggg ggaacccacc 6180  
10 cccacaaaga aagctgaggt cctcctgat gacttgctag ttttaactacc aataacccac 6240  
ttgaattaat catcatcatc aagtcttga taggtgtgag tgggtatcag tggccggtcc 6300  
15 cttcctgggg ctccagcccc cgaggaggcc tcagtgagcc cctgcagaaa atccatgcat 6360  
catgagtgtc tcagggccca gaatatgaga gcaggtagga aacagagaca tcttccatcc 6420  
ctgagaggca gtgcggtcca gtgggtgggg acacgggctc tgggtcaggt ttgtgtgttc 6480  
20 tgtttgtttg ttttgagaca gagtctcgtc ctattgcccc ggctggagtg cagtgtcaca 6540  
atctcggctt actgcaactt ctgccttccc ggattcaagt gattctcctg cctcagcctc 6600  
25 cagagtagct gggattacag gtgcgtgcca ccacgcctgg ctaatttttg tatttttgat 6660  
agagacgggg tttcaccatg ttggccaggc tagtctcgaa ctcttgacct caagtgatct 6720  
gcctgcctcg gcctcccaaa gtgctgggat tacaggcgtg agccaccaca cccagcccca 6780  
30 ggttggtgtt tgaatctgag gagactgaag caccaagggg taaatgttt tgcccacagc 6840  
catacttggg ctcagttcct tgccctacce ctacttgag ctgcttagaa cctggtgggc 6900  
35 acatgggcaa taaccaggtc aactgtttt gtaccaagtg ttatgggaat ccaagatagg 6960

agtaatttgc tctgtggagg ggatgagggg tagtgggttag ggaaagcttc acaaagtggg 7020  
tggttgcttag agattttcca ggtggagaag ggggcttcta ggcagaaggc atagcccaag 7060  
5 caaagactgc aagtgcattg ctgctcatgg gtagaagaga atccaccatt cctcaacatg 7140  
taccgagtcc ttgccatgtg caaggcaaca tgggggtacc aggaattcca agcaatgtcc 7200  
aaacctaggg tctgctttct gggacctgaa gatacaggat ggatcagccc aggctgcaat 7260  
10 cccattacca cgagggggaa aaaaacctga aggctaaatt gtaggtcggg ttagagggtta 7320  
tttatggaaa gttatattct acctacatgg ggtctataag cctggcgcca atcagaaaag 7380  
15 gaacaaacaa cagacctagc tgggagggggc agcattttgt ttaggggggc ggggcacatg 7440  
ttctgggggt acagccagac tcagggtctg tattaatagt ctgagagtaa gacagacaga 7500  
gggatagaag gaaataggtc cctttctctc tctctctctc tctctctctc actctctctc 7560  
20 tctctcacac acacacacag acacacacac acgctctgta ggggtctact tatgctccaa 7620  
gtacaaatca ggccacattt acacaaggag gtaaaggaaa agaacgttgg aggagccaca 7680  
25 ggaccccaaa attccctggt ttccttgaat caggcaggac ttacgcagct gggaggggtg 7740  
agagcctgca gaagccacct gcgagtaagc caagttcaga gtcacagaca caaaagctg 7800  
gtgccatgtc ccacaccgc ccacctccca cctgctcctt gacacagccc tgtgctccac 7860  
30 aaccgggctc ccagatcatt gattatagct ctggggcctg caccgtcctt cctgccacat 7920  
ccccaccca ttcttgaac ctgccctctg tcttctcctt tgtccaaggg caggcaaggg 7980  
35 ctgagctatt gggcagctt gaccaacagc tgaggctcct ttgtggctg gagatgcagg 8040

aggcagggga atattccctct tagtcaatgc gaccatgtgc ctggtttgcc cagggctggc 8100  
tcgtttacac ctgtaggcca agcgtaatta ttaacagctc ccacttctac tctaaaaaat 8160  
5 gacccaatct gggcagtaaa ttatatgggtg cccatgctat taagagctgc aacttgcctgg 8220  
gggtgggtggc tcacacctgt aatcccagta ctttgggacg tcaaggcggg tggatcacct 8280  
gaggtcacga gttagagact ggcttgcca gcatggcaaa accccatctt tactaaaaat 8340  
10 aaaaaatta gcaaggcatg gtggcatgca cctgtaatcc caggtaactcg ggaggctgag 8400  
acaggagaat ggcttgaacc caggaggcag aggttgcagt gagccaagat tgtgccactg 8460  
15 cctccagcc ctggcaacag agcaagactt catctcaaaa gaaaaaggat actgtcaatc 8520  
actgcaggaa gaaccaggt aatgaatgag gagaagagag gggctgagtc accatagtgg 8580  
cagcaccgac tcctgcagga aaggcgagac actgggtcat ggggtactgaa gggtgccctg 8640  
20 aatgacgttc tgctttagag accgaacctg agccctgaaa gtgcatgcct gttcatgggt 8700  
gagagactaa attcatcatt ccttggcagg tactgaatcc tttcttacgg ctgccctcca 8760  
25 atgccaatt tcctacaat tgtctgggt gctaagctt ctgccacca agagggccag 8820  
agctggcagc gacagctgc aggtaggaga gataggtacc cataaggag gtgggaaaga 8880  
gagatggaag gagaggggtg cagagcacac acctcccctg cctgacaact tcctgagggc 8940  
30 tggatcatgcc agcagattta aggcggaggc aggggagatg gggcgggaga ggaagtgaaa 9000  
aaggagaggg tgggatgga gaggaagaga gggatgatcat tcattcattc cattgctact 9060  
35 gactggatgc cagctgtgag ccaggcacca ccctagctct gggcatgtgg ttgtaatctt 9120

ggagcctcat ggagctcaca gggagtgtcg gcaaggagat ggataatgga cggataacaa 9180  
ataaacattt agtacaatgt ccgggaatgg aaagttctcg aaagaaaaat aaagctggtg 9240  
5 agcatataga cagccctgaa ggcggccagg ccaggcattt ctgaggaggt ggcatttgag 9300  
c  
9301

- 10 From the foregoing, it will be appreciated that, although specific embodiments of the invention have been described herein for purposes of illustration, various modifications may be made without deviating from the spirit and scope of the invention. Accordingly, the invention is not limited except as by the appended claims.



CLAIMS

We claim:

1. An isolated nucleic acid molecule selected from the group consisting of:
  - 5 (a) an isolated nucleic acid molecule comprising sequence ID Nos. 1, 5, 9, 11, 13, or, 15, or complementary sequence thereof;
  - (b) an isolated nucleic acid molecule that specifically hybridizes to the nucleic acid molecule of (a) under conditions of high stringency; and
  - (c) an isolated nucleic acid that encodes a TGF-beta binding-protein10 according to (a) or (b).
2. The isolated nucleic acid molecule according to claim 1 wherein said nucleic acid molecule encodes a protein comprising the protein of Sequence ID NO. 2.
3. The isolated nucleic acid molecule according to claim 1 wherein  
15 said nucleic acid molecule encodes a protein comprising the protein of Sequence ID NO. 6.
4. The isolated nucleic acid molecule according to claim 1 wherein said nucleic acid molecule encodes a protein comprising the protein of Sequence ID NO. 10.
- 20 5. The isolated nucleic acid molecule according to claim 1 wherein said nucleic acid molecule encodes a protein comprising the protein of Sequence ID NO. 12.
6. The isolated nucleic acid molecule according to claim 1 wherein  
25 said nucleic acid molecule encodes a protein comprising the protein of Sequence ID NO. 14.
7. The isolated nucleic acid molecule according to claim 1 wherein said nucleic acid molecule encodes a protein comprising the protein of Sequence ID NO. 16.

8. An expression vector, comprising a promoter operably linked to a nucleic acid molecule according to any one of claims 1 to 7.
9. The expression vector according to claim 8 wherein said promoter is selected from the group consisting of CMV I-E promoter, SV40 early promoter and MuLV LTR.
10. The expression vector according to claim 8 wherein said promoter is a tissue-specific promoter.
11. A method of producing a TGF-beta binding protein, comprising, culturing a cell which contains a vector according to claim 8 under conditions and for a time sufficient to produce said protein.
12. The method according to claim 11, further comprising the step of purifying said protein.
13. A viral vector capable of directing the expression of a nucleic acid molecule according to any one of claims 1 to 7.
14. The viral vector according to claim 13 wherein said vector is selected from the group consisting of herpes simplex viral vectors, adenoviral vectors, adenovirus-associated viral vectors and retroviral vectors.
15. A host cell carrying a vector according to any one of claims 8 to 14.
16. The host cell according to claim 15 wherein said cell is selected from the group consisting of a human cell, dog cell, monkey cell, rat cell and mouse cell.
17. An isolated protein, comprising a TGF-beta binding-protein encoded by the nucleic acid molecule according to any one of claims 1 to 7.
18. An antibody which specifically binds to the protein according to claim 17.

19. The antibody according to claim 18 wherein said antibody is a monoclonal antibody.

20. The antibody according to claim 19 wherein said monoclonal antibody is a murine or human antibody.

5 21. The antibody according to claim 18 wherein said antibody is selected from the group consisting of  $F(ab')_2$ ,  $F(ab)_2$ ,  $Fab'$ ,  $Fab$ , and  $Fv$ .

22. A hybridoma which produces an antibody according to claim 19.

23. A fusion protein, comprising a first polypeptide segment comprising a TGF-beta binding-protein encoded by the nucleic acid molecule according to any one of claims 1 to 7, or a portion thereof of at least 10 amino acids in  
10 length, and a second polypeptide segment comprising a non-TGF-beta binding-protein.

24. The fusion protein according to claim 23 wherein said first polypeptide segment is at least 20 amino acids in length.

25. The fusion protein according to claim 23 wherein said first  
15 polypeptide segment is at least 50 amino acids in length.

26. The fusion protein according to claim 23 wherein said second polypeptide comprises multiple anionic amino acid residues.

27. An isolated oligonucleotide which hybridizes to a nucleic acid molecule according to Sequence ID NOs. 1, 3, 5, 7, 9, 11, 13, or 15, or the complement  
20 thereto, under conditions of high stringency.

28. The isolated oligonucleotide according to claim 27 wherein said oligonucleotide is at least 20 nucleotides in length.

29. The isolated oligonucleotide according to claim 27 wherein said oligonucleotide is at least 30 nucleotides in length.

25 30. The isolated oligonucleotide according to claim 27 wherein said

oligonucleotide is at least 50 nucleotides in length.

31. The isolated oligonucleotide according to claim 27 wherein said oligonucleotide is between 50 to 100 nucleotides in length.

32. A pair of primers which specifically amplifies all or a portion of  
5 a nucleic acid molecule according to any one of claims 1 to 7.

33. A ribozyme which cleaves RNA encoding a protein according to claim 17.

34. The ribozyme according to claim 33 wherein said protein comprises the protein of Sequence ID NO. 2.

10 35. The ribozyme according to claim 33 wherein said protein comprises the protein of Sequence ID NO. 6.

36. The ribozyme according to claim 33 wherein said RNA encodes a protein comprising the protein of Sequence ID NO. 10.

15 37. The ribozyme according to claim 33 wherein said RNA encodes a protein comprising the protein of Sequence ID NO. 12.

38. The ribozyme according to claim 33 wherein said RNA encodes a protein comprising the protein of Sequence ID NO. 14.

39. The ribozyme according to claim 33 wherein said RNA encodes a protein comprising the protein of Sequence ID NO. 16.

20 40. The ribozyme according to claim 33 wherein said ribozyme is composed of ribonucleic acids.

41. The ribozyme according to claim 40 wherein one or more of said ribonucleic acids are 2'-O-methyl ribonucleic acids.

42. The ribozyme according to claim 33 wherein said ribozyme is

composed of a mixture of deoxyribonucleic acids and ribonucleic acids.

43. The ribozyme according to claim 33 wherein said ribozyme is composed of nucleic acids having phosphothioate linkages.

44. A nucleic acid molecule comprising a nucleic acid sequence  
5 which encodes a ribozyme according to claim 33.

45. The nucleic acid molecule of claim 44, wherein the nucleic acid is DNA or cDNA.

46. The nucleic acid molecule of claim 44, under the control of a promoter to transcribe the nucleic acid.

10 47. A host cell comprising the ribozyme of claim 33.

48. A vector, comprising the nucleic acid molecule of claim 44.

49. The vector of claim 54, wherein the vector is a plasmid, a virus, retrotransposon or a cosmid.

50. The vector of claim 49 wherein said virus is selected from the  
15 group consisting of retroviruses, adenoviruses, and adeno-associated viruses.

51. A host cell containing the vector according to any one of claims  
48 to 50.

52. The host cell according to claim 51 wherein said host cell is stably transformed with said vector.

20 53. The host cell according to claim 51 wherein the host cell is a human cell.

54. A method for producing a ribozyme, comprising providing DNA encoding the ribozyme according to claim 33 under the transcriptional control of a promoter, and transcribing the DNA to produce the ribozyme.

55. The method of claim 54 wherein the ribozyme is produced *in vitro*.
56. The method of claim 54, further comprising purifying the ribozyme.
- 5 57. A method for increasing bone mineralization, comprising introducing into a warm-blooded animal an effective amount of the ribozyme according to any one of claims 33 to 43.
58. A method of increasing bone mineralization, comprising introducing into a patient an effective amount of the nucleic acid molecule of claim 44,  
10 under conditions favoring transcription of the nucleic acid molecule to produce a ribozyme.
59. A pharmaceutical composition, comprising the ribozyme according to any one of claims 33 to 43, and a pharmaceutically acceptable carrier or diluent.
- 15 60. A pair of primers capable of specifically amplifying all or a portion of a nucleic acid molecule according to any one claims 1 to 7.
61. A method for detecting a nucleic acid molecule which encodes a TGF-beta binding protein, comprising incubating an oligonucleotide according to any one of claims 27 to 31 under conditions of high stringency, and detecting hybridization  
20 of said oligonucleotide.
62. The method according to claim 61 wherein said oligonucleotide is labeled.
63. The method according to claim 61 wherein said oligonucleotide is bound to a solid support.
- 25 64. A method for detecting a TGF-beta binding protein, comprising incubating an antibody according to any one of claims 18 to 21 under conditions and for a time sufficient to permit said antibody to bind to a TGF-beta binding protein, and

detecting said binding.

65. The method according to claim 64 wherein said antibody is bound to a solid support.

66. The method according to claim 64 wherein said antibody is  
5 labeled.

67. The method according to claim 66 wherein said antibody is labeled with a marker selected from the group consisting of enzymes, fluorescent proteins, and radioisotopes.

68. A transgenic animal whose germ cells and somatic cells contain a  
10 nucleic acid molecule encoding a TGF-beta binding-protein according to claim 1 which is operably linked to a promoter effective for the expression of said gene, said gene being introduced into said animal, or an ancestor of said animal, at an embryonic stage, with the proviso that said animal is not a human.

69. The transgenic animal according to claim 68 wherein TGF-beta  
15 binding-protein is expressed from a vector according to any one of claims 8 to 10.

70. A transgenic knockout animal, comprising an animal whose germ cells and somatic cells comprise a disruption of at least one allele of an endogenous nucleic acid molecule which hybridizes to the nucleic acid molecule according to claim 1, wherein said disruption prevents transcription of messenger RNA  
20 from said allele as compared to an animal without said disruption, with the proviso that said animal is not a human.

71. The transgenic animal according to claim 70 wherein said disruption is a nucleic acid deletion, substitution, or, insertion.

72. The transgenic animal according to claim 68 or 70 wherein the  
25 animal is selected from the group consisting of a mouse, a rat and a dog.

73. A method for determining whether a candidate molecule is capable of increasing bone mineral content, comprising:

(a) mixing one or more candidate molecules with TGF-beta-binding-protein encoded by the nucleic acid molecule according to any one of claims 1 to 7 and  
5 a selected member of the TGF-beta family of proteins;

(b) determining whether the candidate molecule alters the signaling of the TGF-beta family member, or alters the binding of the TGF-beta binding-protein to the TGF-beta family member.

74. The method according to claim 73 wherein said member of the  
10 TGF-beta family of proteins is BMP6.

75. A method for determining whether a candidate molecule is capable of increasing bone mineral content, comprising: determining whether a candidate molecule inhibits the binding of TGF-beta binding-protein to bone, or an analogue thereof.

15 76. The method according to claim 75 wherein said analogue of bone is hydroxyapatite.

77. A kit for detection of TGF-beta binding-protein gene expression, comprising a container that comprises a nucleic acid molecule, wherein said nucleic acid molecule is selected from the group consisting of (a) a nucleic acid molecule  
20 comprising the nucleotide sequence of SEQ ID NO: 1, 5, 7, 9, 11, 13, or 15; (b) a nucleic acid molecule comprising the complement of the nucleotide sequence of (a); (c) a nucleic acid molecule that is a fragment of (a) or (b) of at least 20 nucleotides in length.

78. A kit for detection of TGF-beta binding-protein,  
25 comprising a container that comprises an antibody according to any one of claims 18 to 21.

79. An antisense oligonucleotide, comprising a nucleic acid molecule which hybridizes to a nucleic acid molecule according to Sequence ID NOs. 1, 3, 5, 7, 9, 11, 13, or 15, or the complement thereto, and wherein  
30 said oligonucleotide inhibits the expression of TGF-beta binding protein according to



claim 17.

80. The oligonucleotide according to claim 79 wherein said oligonucleotide is 15 nucleotides in length.

81. The oligonucleotide according to claim 79  
5 wherein said oligonucleotide is 20 nucleotides in length.

82. The oligonucleotide according to claim 79 wherein said oligonucleotide is 50 nucleotides in length.

83. The oligonucleotide according to claim 79, wherein said oligonucleotide is comprised of one or more nucleic acid analogs.

84. The oligonucleotide according to claim 79,  
10 wherein said oligonucleotide is comprised of one or more ribonucleic acids.

85. The oligonucleotide according to claim 79, wherein said oligonucleotide is comprised of one or more deoxyribonucleic acids.

86. The oligonucleotide according to claim 79  
15 wherein said oligonucleotide sequence comprises one or more modified covalent linkages.

87. The oligonucleotide according to claim 86 wherein said modified covalent linkage is selected from the group consisting of a phosphorothioate linkage, a phosphotriester linkage, a methyl phosphonate linkage, a  
20 methylene(methylimino) linkage, a morpholino linkage, an amide linkage, a polyamide linkage, a short chain alkyl intersugar linkage, a cycloalkyl intersugar linkage, a short chain heteroatomic intersugar linkage and a heterocyclic intersugar linkage.

1/6

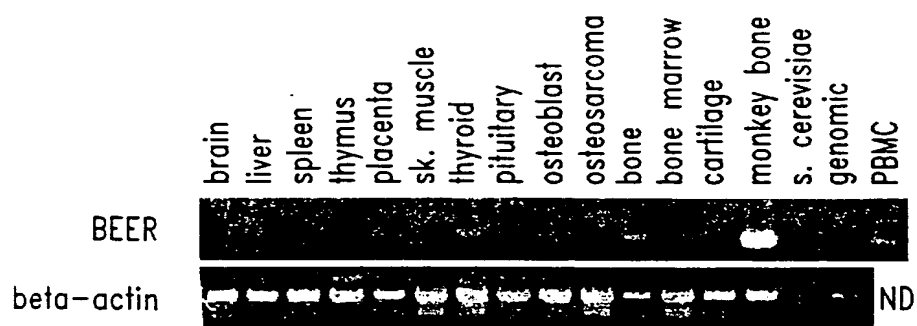
## Common Cysteine Backbone

i		50
human-gremlin.pro	-----	-----
human-cerberus.pro	MHLLFQLLV LLPLGKTRH QGGRNQSSL SPVLLPRNQR ELPTGNHEEA	
human-dan.pro	-----	-----
human-beer.pro	-----	-----
	51	100
human-gremlin.pro	-----M SRTAYTVGAL LLLLGTLPA AEGKKKSOG	
human-cerberus.pro	EKPDLFVAV PHLVAT SPA GEGQRQREKM LSRFGRFWKK PEREMHPSRD	
human-dan.pro	-----	-----
human-beer.pro	-----MQLPLA LCLVCLLVHT	
	101	150
human-gremlin.pro	AI.PPPDKAQ HNDSEQTQSP QQPGSRNRGR GQGRGTAMPG EEVLESSQEA	
human-cerberus.pro	SDSEFPFPGT QSLIQPID G MKMEKSPLRE EAKKFWHHFM FRKTPASQGV	
human-dan.pro	-----MLRVLVGA VL PAMLLAAPP	
human-beer.pro	AFRVVEGQGW QAFKNDATEI IPELGEYPEP PPELENNKTM NRAENGGRPP	
	151	200
human-gremlin.pro	LHVTERKYLK RDWCKTOPLK QTIHEEGCNS RTIINRF.CY GQCNSFYIPR	
human-cerberus.pro	ILPIKSHVH WETCRTPFS QTITHEGCEK VVQNNL.CF GKCGSVHFP.	
human-dan.pro	INKLALFPDK SAWCEAKNIT QIVGHSGCEA KSIQNR.CL GQCFSYSVPN	
human-beer.pro	HHPFETKDVS EYSCRELHFT RYVTDGPCRS AKPVTCLVCS GQCGPARLLP	
	201	250
human-gremlin.pro	HIRKEEGSFQ SCSF...CKP KKFTTMMVTL NCPQLQPPTK K.KRVTRVKQ	
human-cerberus.pro	..GAAQHSHT SCSH...CLP AKFTTMHLPL NCTELSSVIK V...VMLVEE	
human-dan.pro	TFPQSTESLV HCDS...CMP AQSMWEIVTL ECPGHEEVPR VDKLVEKILH	
human-beer.pro	NAIGRGKWWR PSGPDFRCIP DRYRAQRVQL LCPGGEAPRA RKVRLVAS..	
	251	300
human-gremlin.pro	CRC.ISIDLD -----	
human-cerberus.pro	COCKVKTEHE DGHILHAGSQ DSFIPGVSA-----	
human-dan.pro	CSCQACGKEP SHEGLSVYVQ GEDGPGSGPG THPHPHPHPH PGGQTPEPED	
human-beer.pro	CKCKRLTRFH NOSELKDFG EAARPOKGRK PRPRARSAKA NQAELENAY-	
	301	314
human-gremlin.pro	-----	-----
human-cerberus.pro	-----	-----
human-dan.pro	PPGAPHTEEE GAED	
human-beer.pro	-----	-----

Fig. 1

2/6

## Human Beer Gene Expression by RT-PCR

*Fig. 2*

3/6

Fig. 3B

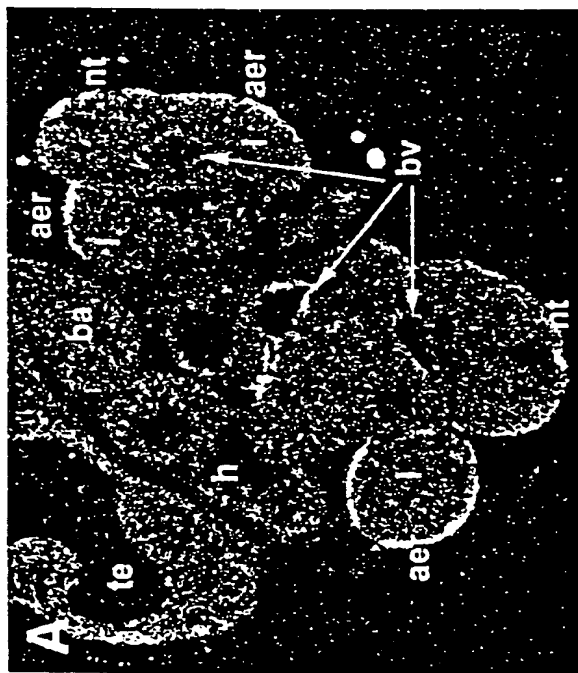
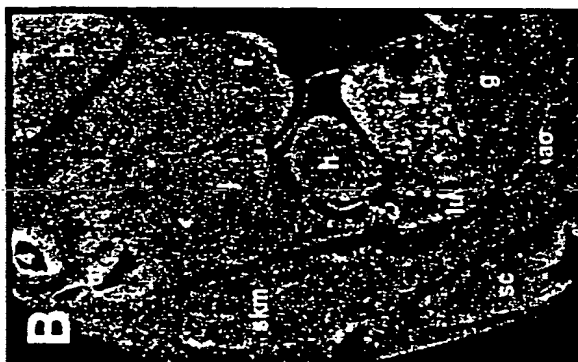


Fig. 3A

Fig. 3D



Fig. 3C

4/6

	Anti Beer		Anti Dan		Anti Gremlin	
<u>BEER Protein</u> (75 ng/lane)	1:5000	1:10,000	1:1000	1:5000	1:1000	1:5000

50 kD —  
34 kD —  
28 kD —  
20 kD —

*Fig. 4A*

	Anti Gremlin		Anti Beer		Anti Dan	
<u>Gremlin Protein</u> (75 ng/lane)	1:5000	1:10,000	1:1000	1:5000	1:1000	1:5000
34 kD —						
28 kD —						
20 kD —						

*Fig. 4B*

	Anti Dan		Anti Beer		Anti Gremlin	
<u>Dan Protein</u> (75 ng/lane)	1:5000	1:10,000	1:1000	1:5000	1:1000	1:5000

34 kD —  
28 kD —  
20 kD —

*Fig. 4C*

5/6

Evaluation of Beer binding to BMP family members  
Anti-FLAG Immunoprecipitation

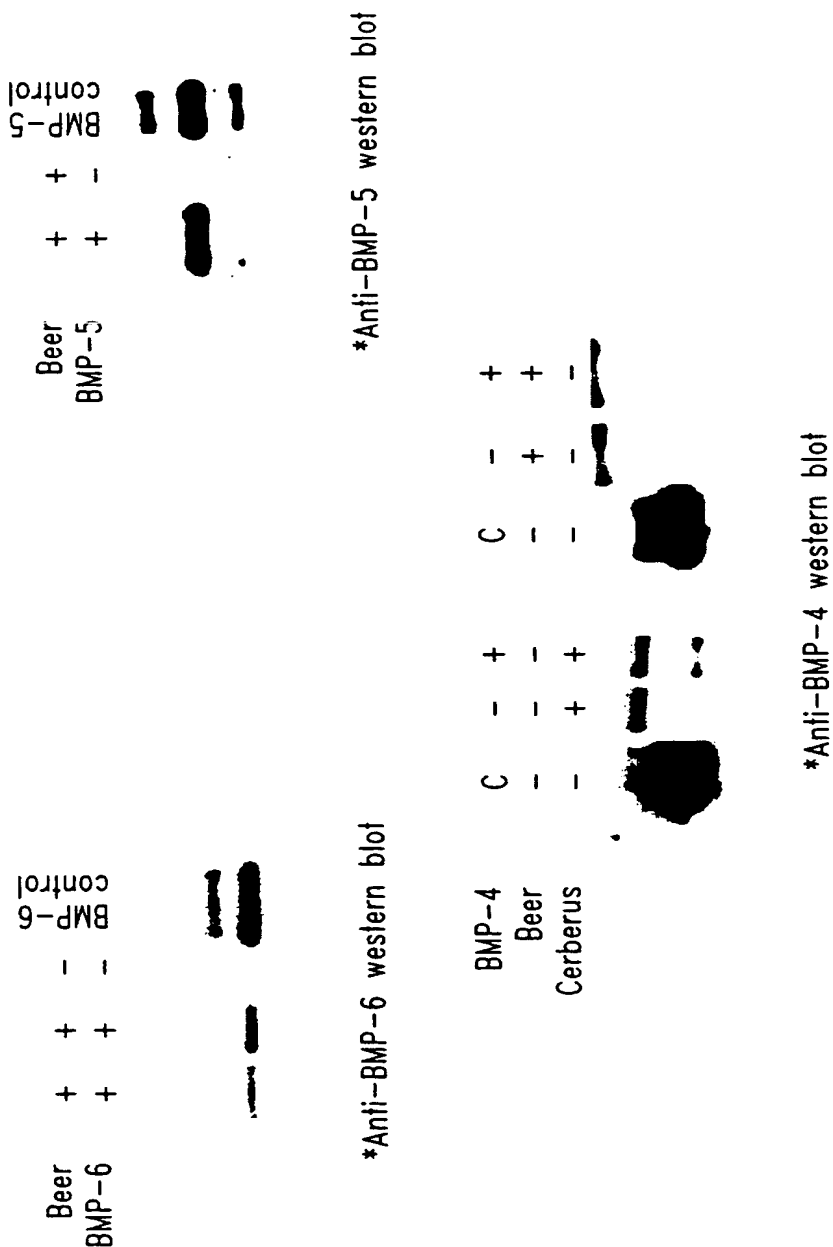
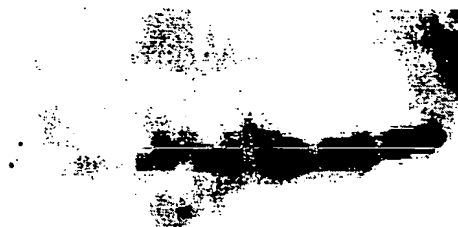


Fig. 5

6 / 6

## BMP-5/Beer Dissociation Constant Characterization

.75 1.5 7.5 15 30 60 120 nM BMP-5



\*Anti-FLAG immunoprecipitation

\*Anti-BMP-5 western blot

## Ionic Disruption of BMP-5/Beer Binding

NaCl(mM)	500	150	150	BMP-5	western	control
Beer	+	+	-			
BMP-5	+	+	+			



\*Anti-FLAG immunoprecipitation

\*Anti-BMP-5 western blot

*Fig. 6*

## SEQUENCE LISTING

5           <110> Brunkow, Mary E.  
            Galas, David J.  
            Kovacevich, Brian  
            Mulligan, John T.  
            Paeper, Bryan W.  
            Van Ness, Jeffrey  
10           Winkler, David G.

15           <120> COMPOSITIONS AND METHODS FOR INCREASING  
            BONE MINERALIZATION

            <130> 240083.508

            <140> US  
            <141> 1999-11-24

20           <160> 41

            <170> FastSEQ for Windows Version 3.0

25           <210> 1  
            <211> 2301  
            <212> DNA  
            <213> Homo sapien.

30           <400> 1

            agagcctgtg ctactggaag gtggcgtgcc ctccctctggc tggtaccatg cagctcccac           60  
            tggccctgtg tctcgtctgc ctgctggtac acacagcctt ccgtgtagtg gagggccagg           120  
            ggtggcaggc gttcaagaat gatgccacgg aaatcatccc cgagctcgga gagtaccccg           180  
            agcctccacc ggagctggag aacaacaaga ccatgaaccg ggcggagAAC ggagggcggc           240  
35           ctccccacca cccctttgag accaaagacg tgtccgagta cagctgccgc gagctgcact           300  
            tcacccgcta cgtgaccgat gggccgtgcc gcagcgccaa gccggtcacc gagctggtgt           360



gctccgggcca gtgcgggcccg gcgcgcctgc tgcaccaacgc catcggccgc ggcaagtgg 420  
 ggcgacctag tgggcccgcac tcccgctgca tccccgaccg ctaccgcgcg cagcgcgtgc 480  
 agctgctgtg tcccgggtggg gaggcgccgc gcgcgcgcaa ggtgcgcctg gtggcctcgt 540  
 gcaagtgcaa gcgcctcacc cgcttccaca accagtcgga gctcaaggac ttcgggaccg 600  
 5 aggcgcgtcg gcgcgagaag ggccggaagc cgcggcccccg cgcgcggagc gccaaagcca 660  
 accaggccga gctggagAAC gcctactaga gcccgcgccg gcccctcccc accggcgggc 720  
 gcccgcggccc tgaacccgcg cccacattt ctgtcctctg cgcgtgggtt gattgtttat 780  
 atttcattgt aaatgcctgc aaccagggc agggggctga gaccttccag gccctgagga 840  
 atcccgggcg ccggcaaggc cccctcagc ccgccagctg aggggtccca cggggcaggg 900  
 10 gaggggaattg agagtcacag acactgagcc acgcagcccc gcctctgggg ccgcctacct 960  
 ttgctgggtc cacttcagag gaggcagaaa tggaaagcatt ttcaccgccc tggggtttta 1020  
 agggagcggg gtgggagtg gaaagtccag ggactgggtt agaaagtgg ataagattcc 1080  
 ccttgcacc tcgctgccc tCagaaagcc tgaggcgtgc ccagagcaca agactggggg 1140  
 caactgtaga tgtggtttct agtcctgggt ctgccactaa cttgctgtgt aacctgaac 1200  
 15 tacacaattc tccctcgga cctcaatttc cactttgtaa aatgagggtg gaggtgggaa 1260  
 taggatctcg aggagactat tggcatatga ttccaaggac tccagtgcct tttgaatggg 1320  
 cagaggtgag agagagagag agaaagagag agaatgaatg cagttgcatt gattcagtgc 1380  
 caaggctact tccagaattc agagttgtga tgctctcttc tgacagccaa agatgaaaaa 1440  
 caaacagaaa aaaaaaagta aagagtctat ttatggctga catatttacg gctgacaaac 1500  
 20 tcctggaaga agctatgtg cttcccagcc tggcttcccc ggatgtttgg ctacctccac 1560  
 cctccatct caaagaaata acatcatcca ttggggtaga aaaggagagg gtccgagggt 1620  
 ggtgggaggg atagaaatca catccgcccc aacttcccaa agagcagcat cctcccccg 1680  
 acccatagcc atgttttaaa gtcaccttcc gaagagaagt gaaaggttca aggacactgg 1740  
 ccttgcaggc ccgagggagc agccatcaca aactcacaga ccagcacatc ccttttgaga 1800  
 25 caccgccttc tgcccaccac tcacggacac atttctgcct agaaaacagc ttcttactgc 1860  
 tcttacatgt gatggcatat cttacactaa aagaatatta ttgggggaaa aactacaagt 1920  
 gctgtacata tgctgagaaa ctgcagagca taatagctgc caccacaaaa tctttttgaa 1980  
 aatcatttcc agacaacctc ttactttctg tgtagttttt aattgttaaa aaaaaaagt 2040  
 tttaaacaga agcacatgac atatgaaagc ctgcaggact ggtcgttttt ttggcaattc 2100  
 30 ttccacgtgg gacttgtcca caagaatgaa agtagtggtt tttaaagagt taagttacat 2160  
 atttattttc tcacttaagt tatttatgca aaagtttttc ttgtagagaa tgacaatgtt 2220  
 aatattgctt tatgaattaa cagtctgttc ttccagagtc cagagacatt gttaataaag 2280  
 acaatgaatc atgaccgaaa g 2301

35 <210> 2

<211> 213

&lt;212&gt; PRT

&lt;213&gt; Homo sapien

&lt;400&gt; 2

5 Met Gln Leu Pro Leu Ala Leu Cys Leu Val Cys Leu Leu Val His Thr  
 1 5 10 15  
 Ala Phe Arg Val Val Glu Gly Gln Gly Trp Gln Ala Phe Lys Asn Asp  
 20 25 30  
 Ala Thr Glu Ile Ile Pro Glu Leu Gly Glu Tyr Pro Glu Pro Pro Pro  
 10 35 40 45  
 Glu Leu Glu Asn Asn Lys Thr Met Asn Arg Ala Glu Asn Gly Gly Arg  
 50 55 60  
 Pro Pro His His Pro Phe Glu Thr Lys Asp Val Ser Glu Tyr Ser Cys  
 65 70 75 80  
 15 Arg Glu Leu His Phe Thr Arg Tyr Val Thr Asp Gly Pro Cys Arg Ser  
 85 90 95  
 Ala Lys Pro Val Thr Glu Leu Val Cys Ser Gly Gln Cys Gly Pro Ala  
 100 105 110  
 Arg Leu Leu Pro Asn Ala Ile Gly Arg Gly Lys Trp Trp Arg Pro Ser  
 20 115 120 125  
 Gly Pro Asp Phe Arg Cys Ile Pro Asp Arg Tyr Arg Ala Gln Arg Val  
 130 135 140  
 Gln Leu Leu Cys Pro Gly Gly Glu Ala Pro Arg Ala Arg Lys Val Arg  
 145 150 155 160  
 25 Leu Val Ala Ser Cys Lys Cys Lys Arg Leu Thr Arg Phe His Asn Gln  
 165 170 175  
 Ser Glu Leu Lys Asp Phe Gly Thr Glu Ala Ala Arg Pro Gln Lys Gly  
 180 185 190  
 Arg Lys Pro Arg Pro Arg Ala Arg Ser Ala Lys Ala Asn Gln Ala Glu  
 30 195 200 205  
 Leu Glu Asn Ala Tyr  
 210

&lt;210&gt; 3

35 &lt;211&gt; 2301

&lt;212&gt; DNA

&lt;213&gt; Homo sapien

&lt;400&gt; 3

	agagcctgtg ctactggaag gtggcgtgcc ctccctctggc tggtagcatg cagctcccccac	60
5	tggccctgtg tctcgtctgc ctgctggtac acacagcctt ccgtgtagtg gagggctagg	120
	ggtggcaggg gtccaagaat gatgccacgg aaatcatccc cgagctcgga gaggaccccg	180
	agcctccacc ggagctggag aacaacaaga ccatgaaccg ggccggagaac ggagggcggc	240
	ctccccacca cccctttgag accaaagacg tgtccgagta cagctgccgc gagctgcact	300
	tcacccgcta cgtgaccgat gggccgtgcc gcagcgccaa gccggtcacc gagctgggtgt	360
10	gctccggcca gtgcggccccg gcgcgcctgc tggccaacgc catcggccgc ggcaagtggc	420
	ggcgacctag tgggccccgac ttccgctgca tccccgaccg ctaccgcgcg cagcgcgctgc	480
	agctgctgtg tccccgggtggg gagggcgccgc gcgcgcgcaa ggtgcgcctg gtggcctcgt	540
	gcaagtgcga gcgcctcacc cgcttccaca accagtcgga gctcaaggac ttcgggaccg	600
	aggccgctcg gccgcagaag ggccggaagc cgccggccccg cgcccgagag gccaaagcca	660
15	accaggccga gctggagaac gcctactaga gcccgccccg gccctcccc accggcgggc	720
	gccccggccc tgaaccgcgc ccccacattt ctgtccctctg cgcgtgggttt gattgtttat	780
	atcttcattgt aaatgcctgc aaccaggggc agggggctga gaccttccag gccctgagga	840
	atccccgggcg ccggcaaggc cccctcagc ccgccagctg aggggtccca cggggcaggg	900
	gaggggaattg agagtcacag aactgagcc acgcagcccc gcctctgggg ccgcctacct	960
20	ttgctgggtcc cacttcagag gaggcagaaa tggaaagcatt ttcaccgccc tggggtttta	1020
	agggagcggg gtgggagtg gaaagtcag ggactgggtta agaaagtgtg ataagattcc	1080
	cccttgccacc tcgctgcccc tcagaaagcc tgaggcgtgc ccagagcaca agactggggg	1140
	caactgtaga tgtgggtttt agtcctggct ctgccactaa cttgctgtgt aaccttgaac	1200
	tacacaattc tcttcggga cctcaatttc cactttgttaaatgaggggtg gaggtgggaa	1260
25	taggatctcg aggagactat tggcatatga ttccaaggac tccagtgcct tttgaatggg	1320
	cagagggtgag agagagagag agaaagagag agaataatg cagttgcatt gattcagtgc	1380
	caaggctcact tccagaattc agagtgtgta tgctctcttc tgacagccaa agatgaaaaa	1440
	caaacagaaa aaaaaaagta aagagtctat ttatggctga catatttacg gctgacaaac	1500
	tcctggaaga agctatgctg ctccccagcc tggcttcccc ggatgtttgg ctacctccac	1560
30	ccctccatct caaagaaata acatcatcca ttggggtaga aaaggagagg gtccgagggg	1620
	gggtgggaggg atagaaatca catccgcccc aacttcccaa agagcagcat cccctccccg	1680
	acccatagcc atgttttaaa gtcaccttcc gaagagaagt gaaagggtca aggacactgg	1740
	cccttgaggc ccgagggagc agccatcaca aactcacaga ccagcacatc ccttttgaga	1800
	caccgccttc tgcccaccac tcacggacac atttctgcct agaaaacagc ttcttactgc	1860
35	tcttacatgt gatggcatat cttacactaa aagaatatta ttgggggaaa aactacaagt	1920
	gctgtacata tgctgagaaa ctgcagagca taatagctgc cacccaaaaa tctttttgaa	1980

aatcatttcc agacaacctc ttactttctg tgtagtcttt aattgtttaaa aaaaaaaagt 2040  
 tttaaacaga agcatatgac atatgaaagc ctgcaggact ggctgttttt ttggcaattc 2100  
 tttcacgtgg gacttgtcca caagaatgaa agtagtggtt tttaaagagt taagttacat 2160  
 atttattttc tcacttaagt tatttatgca aaagtctttc ttgtagagaa tgacaatgtt 2220  
 5 aatattgctt tatgaattaa cagtctgttc tttcagagtc cagagacatt gttaataaag 2280  
 acaatgaatc atgaccgaaa g 2301

<210> 4

<211> 23

10 <212> PRT

<213> Homo sapien

<400> 4

Met Gln Leu Pro Leu Ala Leu Cys Leu Val Cys Leu Leu Val His Thr  
 15 1 5 10 15  
 Ala Phe Arg Val Val Glu Gly  
 20

<210> 5

20 <211> 2301

<212> DNA

<213> Homo sapien

<400> 5

25 agagcctgtg ctactggaag gtggcgtgcc ctctctctggc tgggtaccatg cagctcccccac 60  
 tggccctgtg tctcatctgc ctgctggtac acacagcctt ccgtgtagtg gagggccagg 120  
 ggtggcaggc gttcaagaat gatgccacgg aaatcatccg cgagctcgga gagtaccccg 180  
 agcctccacc ggagctggag aacaacaaga ccatgaaccg ggcggagaac ggagggcggc 240  
 ctccccacca cccctttgag accaaagacg tgtccgagta cagctgccgc gagctgcact 300  
 30 tcaccgcta cgtgaccgat gggccgtgcc gcagcgccaa gccggtcacc gagctggtgt 360  
 gctccggcca gtgcggcccg gcgcgcctgc tgcccaacgc catcggccgc ggcaagtggc 420  
 ggcgacctag tgggcccgcac tccgctgca tccccgaccg ctaccgcgcg cagcgcgtgc 480  
 agctgctgtg tcccgggtggg gaggcgcgcg gcgcgcgcaa ggtgcgcctg gtggcctcgt 540  
 gcaagtgcaa gcgcctcacc cgcttcacaa accagtcgga gctcaaggac ttcgggaccg 600  
 35 aggccgctcg gccgcagaag ggccggaagc cgcggccccc cgcccggagc gccaaagcca 660  
 accaggccga gctggagaac gcctactaga gcccgcgcgc gccctcccc accggcgggc 720

gccccggccc tgaacccgcg cccacattt ctgtccctctg cgcgtgggtt gattgtttat 780  
 atttcattgt aaatgcctgc aaccagggc agggggctga gacctccag gccctgagga 840  
 atccccggcg ccggcaaggc cccctcagc ccgccagctg aggggtccca cggggcaggg 900  
 gaggggaattg agagtcacag aactgagcc acgcagcccc gcctctgggg cgcctacct 960  
 5 ttgctgggtcc cacttcagag gaggcagaaa tggaaagcatt ttcaccgccc tgggggtttta 1020  
 agggagcggg gtgggagtg gaaagtcag ggactgggta agaaagtgg ataagattcc 1080  
 cccctgcacc tcgctgcccc tcagaaagcc tgaggcgtgc ccagagcaca agactggggg 1140  
 caactgtaga tgtgggtttct agtcctggct ctgccactaa ctgctgtgt aacctgaac 1200  
 tacacaattc tccttcggga cctcaatttc cactttgtaa aatgaggggt gaggtgggaa 1260  
 10 taggatctcg aggagactat tggcatatga tccaaggac tccagtgcct ttggaatggg 1320  
 cagaggtgag agagagagag agaaagagag agaataaatg cagtgcatt gattcagtgc 1380  
 caaggtcact tccagaattc aagattgtga tgctctcttc tgacagccaa agatgaaaaa 1440  
 caaacagaaa aaaaaaagta aagattctat ttatggctga catatttacg gctgacaaac 1500  
 tcctggaaga agctatgctg ctcccagcc tggcttcccc ggatgtttgg ctacctccac 1560  
 15 cccctccatct caaagaaata acatcatcca ttggggtaga aaaggagagg gtccgagggt 1620  
 ggtgggaggg atagaaatca catccgcccc aacttcccaa agagcagcat cctcccccg 1680  
 acccatagcc atgttttaaa gtcaccttcc gaagagaagt gaaagggtca aggacactgg 1740  
 ccttgcaggc ccgagggagc agccatcaca aactcacaga ccagcacatc ccttttgaga 1800  
 caccgccttc tgcccaccac tcacggacac attcttgcct agaaaacagc ttcttactgc 1860  
 20 tcttacatgt gatggcatat cttacactaa aagaatatta ttgggggaaa aactacaagt 1920  
 gctgtacata tgctgagaaa ctgcagagca taatagctgc cacccaaaaa tctttttgaa 1980  
 aatcatttcc agacaacctc ttactttctg tgtagttttt aattgttaaa aaaaaaagt 2040  
 tttaaacaga agcacatgac atatgaaagc ctgcaggact ggtcgttttt ttggcaattc 2100  
 ttccacgtgg gacttgtcca caagaatgaa agtagtgggt tttaaagagt taagttacat 2160  
 25 atttattttc tcacttaagt tatttatgca aaagtttttc ttgtagagaa tgacaatgtt 2220  
 aatattgctt tatgaattaa cagtctgttc ttccagagtc cagagacatt gttataaag 2280  
 acaatgaatc atgaccgaaa g 2301

&lt;210&gt; 6

30

&lt;211&gt; 213

&lt;212&gt; PRT

&lt;213&gt; Homo sapien

&lt;400&gt; 6

35

Met Gln Leu Pro Leu Ala Leu Cys Leu Ile Cys Leu Leu Val His Thr

1

5

10

15

Ala Phe Arg Val Val Glu Gly Gln Gly Trp Gln Ala Phe Lys Asn Asp  
 20 25 30  
 Ala Thr Glu Ile Ile Arg Glu Leu Gly Glu Tyr Pro Glu Pro Pro Pro  
 35 40 45  
 5 Glu Leu Glu Asn Asn Lys Thr Met Asn Arg Ala Glu Asn Gly Gly Arg  
 50 55 60  
 Pro Pro His His Pro Phe Glu Thr Lys Asp Val Ser Glu Tyr Ser Cys  
 65 70 75 80  
 Arg Glu Leu His Phe Thr Arg Tyr Val Thr Asp Gly Pro Cys Arg Ser  
 85 90 95  
 10 Ala Lys Pro Val Thr Glu Leu Val Cys Ser Gly Gln Cys Gly Pro Ala  
 100 105 110  
 Arg Leu Leu Pro Asn Ala Ile Gly Arg Gly Lys Trp Trp Arg Pro Ser  
 115 120 125  
 15 Gly Pro Asp Phe Arg Cys Ile Pro Asp Arg Tyr Arg Ala Gln Arg Val  
 130 135 140  
 Gln Leu Leu Cys Pro Gly Gly Glu Ala Pro Arg Ala Arg Lys Val Arg  
 145 150 155 160  
 Leu Val Ala Ser Cys Lys Cys Lys Arg Leu Thr Arg Phe His Asn Gln  
 165 170 175  
 20 Ser Glu Leu Lys Asp Phe Gly Thr Glu Ala Ala Arg Pro Gln Lys Gly  
 180 185 190  
 Arg Lys Pro Arg Pro Arg Ala Arg Ser Ala Lys Ala Asn Gln Ala Glu  
 195 200 205  
 25 Leu Glu Asn Ala Tyr  
 210

&lt;210&gt; 7

&lt;211&gt; 2301

30

&lt;212&gt; DNA

&lt;213&gt; Homo sapien

&lt;400&gt; 7

agagcctgtg ctactggaag gtggcggtgcc ctccctctggc tggtagcatg cagctccac 60  
 35 tggccctgtg tctcgtctgc ctgctggtac acacagcctt ccgtgtagtg gagggccagg 120  
 ggtggcaggc gttcaagaat gatgccacgg aaatcatccg cgagctcgga gactacccc 180

	agcctccacc	ggagctggag	aacaacaaga	ccatgaaccg	ggcggagaac	ggagggcggc	240
	ctccccacca	cccccttgag	accaaagacg	tgtccgagta	cagctgccgc	gagctgcact	300
	tcacccgcta	cgtgaccgat	gggcccgtgc	gcagcgccaa	gccggtcacc	gagctggtgt	360
	gctccggcca	gtgcggcccc	gcgcgcctgc	tgcccaacgc	catcgccgc	ggcaagtgg	420
5	ggcgacctag	tgggcccagac	ttccgctgca	tccccgaccg	ctaccgcgcg	cagcgctgc	480
	agctgctgtg	tcccgggtgg	gaggcgccgc	gcgcgcgcaa	ggtgcgcctg	gtggcctcgt	540
	gcaagtgcga	gcgcctcacc	cgcttcacac	accagtcgga	gctcaaggac	ttcgggaccg	600
	aggccgctcg	gccgcagaag	ggccggaagc	cgcggccccc	cgccccggagc	gccaaagcca	660
	accaggccga	gctggagAAC	gcctactaga	gcccgcgccgc	gccccctccc	accggcgggc	720
10	gccccggccc	tgaacccgcg	ccccacattt	ctgtcctctg	cgcgtggttt	gattgtttat	780
	atttcattgt	aaatgcctgc	aaccagggc	agggggctga	gaccttccag	gccctgagga	840
	atccccggcg	ccggcaaggc	ccccctcagc	ccgccagctg	aggggtccca	cggggcaggg	900
	gagggaaattg	agagtcacag	acactgagcc	acgcagcccc	gcctctgggg	ccgcctacct	960
	ttgctggctc	cacttcagag	gaggcagaaa	tggaaagcatt	ttcaccgccc	tggggtttta	1020
15	agggagcgg	gtgggagtg	gaaagtccag	ggactgggtta	agaaagtgg	ataagattcc	1080
	cccttgacc	tcgctgccc	tcagaaagcc	tggggcgtgc	ccagagcaca	agactggggg	1140
	caactgtaga	tgtggtttct	agtcctggct	ctgccactaa	cttgctgtgt	aaccttgaac	1200
	tacacaattc	tccttcggga	cctcaatttc	cactttgtaa	aatgaggggtg	gaggtgggaa	1260
	taggatctcg	aggagactat	tggcatatga	ttccaaggac	tccagtgcct	tttgaatggg	1320
20	cagaggtgag	agagagagag	agaaagagag	agaatgaatg	cagttgcatt	gattcagtgc	1380
	caaggtcact	tcagaattc	agagttgtga	tgtctctctc	tgacagccaa	agatgaaaaa	1440
	caaacagaaa	aaaaaaagta	aagagtctat	ttatggctga	catattttacg	gctgacaaac	1500
	tcctggaaga	agctatgctg	cttcccagcc	tggcttcccc	ggatgtttgg	ctacctccac	1560
	ccctccatct	caaagaaata	acatcatcca	ttggggtaga	aaaggagagg	gtccgaggg	1620
25	gggtgggagg	atagaaatca	catccgcccc	aacttcccaa	agagcagcat	ccctcccccg	1680
	acccatagcc	atgttttaaa	gtcaccttcc	gaagagaagt	gaaaggttca	aggacactgg	1740
	ccttgaggc	ccgagggagc	agccatcaca	aactcacaga	ccagcacatc	ccttttgaga	1800
	caccgccttc	tgcccaccac	tcacggacac	atttctgcct	agaaaacagc	ttcttactgc	1860
	tcttacatgt	gatggcatat	cttacctaa	agaatatta	ttgggggaaa	aactacaagt	1920
30	gctgtacata	tgttgagaaa	ctgcagagca	taatagctgc	cacccaaaaa	tctttttgaa	1980
	aatcatttcc	agacaacctc	ttactttctg	tgtagttttt	aattgttaaa	aaaaaaaagt	2040
	tttaaacaga	agcacatgac	atatgaaagc	ctgcaggact	ggtcgttttt	ttggcaattc	2100
	ttccacgtgg	gacttgtcca	caagaatgaa	agtagtggtt	tttaaagagt	taagttacat	2160
	atatttttcc	tcacttaagt	tatttatgca	aaagtttttc	ttgtagagaa	tgacaatgtt	2220
35	aatattgctt	tatgaattaa	cagtctgttc	ttccagagtc	cagagacatt	gttaataaag	2280
	acaatgaatc	atgaccgaaa	g				2301

&lt;210&gt; 8

&lt;211&gt; 213

&lt;212&gt; PRT

5 &lt;213&gt; Homo sapien

&lt;400&gt; 8

Met Gln Leu Pro Leu Ala Leu Cys Leu Val Cys Leu Leu Val His Thr  
 1 5 10 15  
 10 Ala Phe Arg Val Val Glu Gly Gln Gly Trp Gln Ala Phe Lys Asn Asp  
 20 25 30  
 Ala Thr Glu Ile Ile Arg Glu Leu Gly Glu Tyr Pro Glu Pro Pro Pro  
 35 40 45  
 Glu Leu Glu Asn Asn Lys Thr Met Asn Arg Ala Glu Asn Gly Gly Arg  
 15 50 55 60  
 Pro Pro His His Pro Phe Glu Thr Lys Asp Val Ser Glu Tyr Ser Cys  
 65 70 75 80  
 Arg Glu Leu His Phe Thr Arg Tyr Val Thr Asp Gly Pro Cys Arg Ser  
 85 90 95  
 20 Ala Lys Pro Val Thr Glu Leu Val Cys Ser Gly Gln Cys Gly Pro Ala  
 100 105 110  
 Arg Leu Leu Pro Asn Ala Ile Gly Arg Gly Lys Trp Trp Arg Pro Ser  
 115 120 125  
 Gly Pro Asp Phe Arg Cys Ile Pro Asp Arg Tyr Arg Ala Gln Arg Val  
 25 130 135 140  
 Gln Leu Leu Cys Pro Gly Gly Glu Ala Pro Arg Ala Arg Lys Val Arg  
 145 150 155 160  
 Leu Val Ala Ser Cys Lys Cys Lys Arg Leu Thr Arg Phe His Asn Gln  
 165 170 175  
 30 Ser Glu Leu Lys Asp Phe Gly Thr Glu Ala Ala Arg Pro Gln Lys Gly  
 180 185 190  
 Arg Lys Pro Arg Pro Arg Ala Arg Ser Ala Lys Ala Asn Gln Ala Glu  
 195 200 205  
 Leu Glu Asn Ala Tyr  
 35 210



<210> 9  
 <211> 642  
 <212> DNA  
 <213> Cercopithecus pygerythrus

5

&lt;400&gt; 9

```

atgcagctcc cactggccct gtgtcttgtc tgctgctgg tacacgcagc ctcccggtga      60
gtggaggggcc aggggtggca ggccttcaag aatgatgcca cggaaatcat ccccgagctc     120
ggagagtacc ccgagcctcc accggagctg gagaacaaca agaccatgaa ccgggaggag      180
10 aatggagggc ggccctcccca ccaccccttt gagaccaaag acgtgtccga gtacagctgc     240
cgagagctgc atttcacccg ctacgtgacc gatggggccgt gccgcagcgc caagccagtc     300
accgagtttg tgtgctccgg ccagtgcggc ccggcacgcc tgctgcccga cgccatcggc     360
cgcgggcaagt ggtggcgccc gaggggccc gacttccgct gcatccccga ccgctaccgc     420
gcgcagcgtg tgcagctgct gtgtcccggt ggtgccgcgc cgcgcgcgcg caaggtgcgc     480
15 ctggtggcct cgtgcaagtg caagcgctc acccgcttcc acaaccagtc ggagctcaag     540
gacttcgggtc ccgaggccgc tcggccgcag aagggccgga agccgcggcc ccgcgcccgg     600
ggggccaaag ccaatcaggc cgagctggag aacgcctact ag                          642
  
```

<210> 10  
 20 <211> 213  
 <212> PRT  
 <213> Cercopithecus pygerythrus

&lt;400&gt; 10

```

25 Met Gln Leu Pro Leu Ala Leu Cys Leu Val Cys Leu Leu Val His Ala
    1           5           10           15
Ala Phe Arg Val Val Glu Gly Gln Gly Trp Gln Ala Phe Lys Asn Asp
    20           25           30
Ala Thr Glu Ile Ile Pro Glu Leu Gly Glu Tyr Pro Glu Pro Pro Pro
30           35           40           45
Glu Leu Glu Asn Asn Lys Thr Met Asn Arg Ala Glu Asn Gly Gly Arg
    50           55           60
Pro Pro His His Pro Phe Glu Thr Lys Asp Val Ser Glu Tyr Ser Cys
    65           70           75           80
35 Arg Glu Leu His Phe Thr Arg Tyr Val Thr Asp Gly Pro Cys Arg Ser
    85           90           95
  
```

Ala Lys Pro Val Thr Glu Leu Val Cys Ser Gly Gln Cys Gly Pro Ala  
 100 105 110  
 Arg Leu Leu Pro Asn Ala Ile Gly Arg Gly Lys Trp Trp Arg Pro Ser  
 115 120 125  
 5 Gly Pro Asp Phe Arg Cys Ile Pro Asp Arg Tyr Arg Ala Gln Arg Val  
 130 135 140  
 Gln Leu Leu Cys Pro Gly Gly Ala Ala Pro Arg Ala Arg Lys Val Arg  
 145 150 155 160  
 Leu Val Ala Ser Cys Lys Cys Lys Arg Leu Thr Arg Phe His Asn Gln  
 10 165 170 175  
 Ser Glu Leu Lys Asp Phe Gly Pro Glu Ala Ala Arg Pro Gln Lys Gly  
 180 185 190  
 Arg Lys Pro Arg Pro Arg Ala Arg Gly Ala Lys Ala Asn Gln Ala Glu  
 195 200 205  
 15 Leu Glu Asn Ala Tyr  
 210

&lt;210&gt; 11

&lt;211&gt; 638

20 &lt;212&gt; DNA

&lt;213&gt; Mus musculus

&lt;400&gt; 11

atgcagccct cactagcccc gtgcctcatc tgccctacttg tgcacgctgc cttctgtgct 60  
 25 gtggagggcc aggggtggca agccttcagg aatgatgcc aagaggtcat cccagggctt 120  
 ggagagtacc ccgagcctcc tcttgagaac aaccagacca tgaaccgggc ggagaatgga 180  
 ggagagacct cccaccatcc ctatgacgcc aaaggtgtgt ccgagtacag ctgccgcgag 240  
 ctgcactaca cccgcttcc gacagacggc ccatgccgca gcgccaagcc ggtcaccgag 300  
 ttggtgtgct ccggccagtg cggccccgcg cggctgctgc ccaacgccat cgggcgctg 360  
 30 aagtgggtggc gcccgaacgg accggatttc cgctgcatcc cggatcgcta ccgcgcgag 420  
 cgggtgcagc tgctgtgccc cgggggcgcg gcgcccgcgt cgcgcaaggt gcgtctggtg 480  
 gcctcgtgca agtgcaagcg cctcaccgcg tccacaacc agtcggagct caaggacttc 540  
 gggccggaga ccgcgcggcc gcagaagggt cgcaagccgc ggcccggcgc ccggggagcc 600  
 aaagccaacc aggcggagct ggagaacgcc tactagag 638  
 35

&lt;210&gt; 12

&lt;211&gt; 211

&lt;212&gt; PRT

&lt;213&gt; Mus musculus

5 &lt;400&gt; 12

Met Gln Pro Ser Leu Ala Pro Cys Leu Ile Cys Leu Leu Val His Ala  
 1 5 10 15  
 Ala Phe Cys Ala Val Glu Gly Gln Gly Trp Gln Ala Phe Arg Asn Asp  
 20 25 30  
 10 Ala Thr Glu Val Ile Pro Gly Leu Gly Glu Tyr Pro Glu Pro Pro Pro  
 35 40 45  
 Glu Asn Asn Gln Thr Met Asn Arg Ala Glu Asn Gly Gly Arg Pro Pro  
 50 55 60  
 His His Pro Tyr Asp Ala Lys Asp Val Ser Glu Tyr Ser Cys Arg Glu  
 15 65 70 75 80  
 Leu His Tyr Thr Arg Phe Leu Thr Asp Gly Pro Cys Arg Ser Ala Lys  
 85 90 95  
 Pro Val Thr Glu Leu Val Cys Ser Gly Gln Cys Gly Pro Ala Arg Leu  
 100 105 110  
 20 Leu Pro Asn Ala Ile Gly Arg Val Lys Trp Trp Arg Pro Asn Gly Pro  
 115 120 125  
 Asp Phe Arg Cys Ile Pro Asp Arg Tyr Arg Ala Gln Arg Val Gln Leu  
 130 135 140  
 Leu Cys Pro Gly Gly Ala Ala Pro Arg Ser Arg Lys Val Arg Leu Val  
 25 145 150 155 160  
 Ala Ser Cys Lys Cys Lys Arg Leu Thr Arg Phe His Asn Gln Ser Glu  
 165 170 175  
 Leu Lys Asp Phe Gly Pro Glu Thr Ala Arg Pro Gln Lys Gly Arg Lys  
 180 185 190  
 30 Pro Arg Pro Gly Ala Arg Gly Ala Lys Ala Asn Gln Ala Glu Leu Glu  
 195 200 205  
 Asn Ala Tyr  
 210

35 &lt;210&gt; 13

&lt;211&gt; 674

&lt;212&gt; DNA

&lt;213&gt; Rattus norvegicus

&lt;400&gt; 13

5 gaggaccgag tgccttccct ccttctggca ccatgcagct ctactagcc ccttgcccttg 60  
 cctgcctgct tgtacatgca gccttcgttg ctgtggagag ccaggggtgg caagccttca 120  
 agaatgatgc cacagaaatc atcccgggac tcagagagta cccagagcct cctcaggaac 180  
 tagagaacaa ccagaccatg aaccggggccg agaacggagg cagaccccc caccatcctt 240  
 atgacaccaa agacgtgtcc gactacagct gccgcgagct gcaactacacc cgcttcgtga 300  
 10 ccgacggccc gtgcgcagc gccaaagccgg tcaccgagtt ggtgtgctcg ggccagtgcg 360  
 gccccgcgcg gctgctgccc aacgccatcg ggcgcgtgaa gtggtggcgc ccgaacggac 420  
 ccgacttccg ctgcatcccg gatcgctacc gcgcgcagcg ggtgcagctg ctgtgccccg 480  
 gcggcgcggc gccgcgctcg cgcaagggtgc gtctggtggc ctctgcaag tgcaagcgcc 540  
 tcaccgcctt ccacaaccag tcggagctca aggacttcgg acctgagacc gcgcggccgc 600  
 15 agaagggtcg caagccgcgg ccccgcgccc ggggagccaa agccaaccag gcggagctgg 660  
 agaacgccta ctag 674

&lt;210&gt; 14

&lt;211&gt; 213

20

&lt;212&gt; PRT

&lt;213&gt; Rattus norvegicus

&lt;400&gt; 14

Met Gln Leu Ser Leu Ala Pro Cys Leu Ala Cys Leu Leu Val His Ala  
 25 1 5 10 15  
 Ala Phe Val Ala Val Glu Ser Gln Gly Trp Gln Ala Phe Lys Asn Asp  
 20 25 30  
 Ala Thr Glu Ile Ile Pro Gly Leu Arg Glu Tyr Pro Glu Pro Pro Gln  
 35 40 45  
 30 Glu Leu Glu Asn Asn Gln Thr Met Asn Arg Ala Glu Asn Gly Gly Arg  
 50 55 60  
 Pro Pro His His Pro Tyr Asp Thr Lys Asp Val Ser Glu Tyr Ser Cys  
 65 70 75 80  
 Arg Glu Leu His Tyr Thr Arg Phe Val Thr Asp Gly Pro Cys Arg Ser  
 35 85 90 95  
 Ala Lys Pro Val Thr Glu Leu Val Cys Ser Gly Gln Cys Gly Pro Ala

100 105 110  
 Arg Leu Leu Pro Asn Ala Ile Gly Arg Val Lys Trp Trp Arg Pro Asn  
 115 120 125  
 Gly Pro Asp Phe Arg Cys Ile Pro Asp Arg Tyr Arg Ala Gln Arg Val  
 5 130 135 140  
 Gln Leu Leu Cys Pro Gly Gly Ala Ala Pro Arg Ser Arg Lys Val Arg  
 145 150 155 160  
 Leu Val Ala Ser Cys Lys Cys Lys Arg Leu Thr Arg Phe His Asn Gln  
 165 170 175  
 10 Ser Glu Leu Lys Asp Phe Gly Pro Glu Thr Ala Arg Pro Gln Lys Gly  
 180 185 190  
 Arg Lys Pro Arg Pro Arg Ala Arg Gly Ala Lys Ala Asn Gln Ala Glu  
 195 200 205  
 Leu Glu Asn Ala Tyr  
 15 210  
  
 <210> 15  
 <211> 532  
 <212> DNA  
 20 <213> Bos torus  
  
 <400> 15  
 agaatgatgc cacagaaatc atccccgagc tgggcgagta ccccgagcct ctgccagagc 60  
 tgaacaacaa gaccatgaac cgggcgagga acggagggag acctccccac caccctttg 120  
 25 agaccaaaga cgcctccgag tacagctgcc gggagctgca cttcaccgcg tacgtgaccg 180  
 atgggcccgtg ccgcagcgcc aagccgggtca ccgagctggt gtgctcgggc cagtgcggcc 240  
 cggcgcgccct gctgcccaac gccatcggcc gcggcaagtg gtggcgccca agcgggccccg 300  
 acttccgctg catccccgac cgctaccgcg cgcagcgggt gcagctgttg tgcctggcg 360  
 gcgcggcgcc gcgcgcgcgc aaggtgcgcc tgggtggcctc gtgcaagtgc aagcgccctca 420  
 30 ctgcgttcca caaccagtcc gagctcaagg acttcggggc cgaggccgcg cggccgcaaa 480  
 cgggcccggaa gctgcggccc cgcgcccggg gcaccaaagc cagccgggccc ga 532  
  
 <210> 16  
 <211> 176  
 35 <212> PRT  
 <213> Bos torus

15

&lt;400&gt; 16

Asn Asp Ala Thr Glu Ile Ile Pro Glu Leu Gly Glu Tyr Pro Glu Pro  
 1 5 10 15  
 5 Leu Pro Glu Leu Asn Asn Lys Thr Met Asn Arg Ala Glu Asn Gly Gly  
 20 25 30  
 Arg Pro Pro His His Pro Phe Glu Thr Lys Asp Ala Ser Glu Tyr Ser  
 35 40 45  
 Cys Arg Glu Leu His Phe Thr Arg Tyr Val Thr Asp Gly Pro Cys Arg  
 10 50 55 60  
 Ser Ala Lys Pro Val Thr Glu Leu Val Cys Ser Gly Gln Cys Gly Pro  
 65 70 75 80  
 Ala Arg Leu Leu Pro Asn Ala Ile Gly Arg Gly Lys Trp Trp Arg Pro  
 85 90 95  
 15 Ser Gly Pro Asp Phe Arg Cys Ile Pro Asp Arg Tyr Arg Ala Gln Arg  
 100 105 110  
 Val Gln Leu Leu Cys Pro Gly Gly Ala Ala Pro Arg Ala Arg Lys Val  
 115 120 125  
 Arg Leu Val Ala Ser Cys Lys Cys Lys Arg Leu Thr Arg Phe His Asn  
 20 130 135 140  
 Gln Ser Glu Leu Lys Asp Phe Gly Pro Glu Ala Ala Arg Pro Gln Thr  
 145 150 155 160  
 Gly Arg Lys Leu Arg Pro Arg Ala Arg Gly Thr Lys Ala Ser Arg Ala  
 165 170 175

25

&lt;210&gt; 17

&lt;211&gt; 35828

&lt;212&gt; DNA

&lt;213&gt; Mus musculus

30

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;222&gt; (1)...(35828)

&lt;223&gt; n = A,T,C or G

35

&lt;400&gt; 17

	cgcggttttgg	tgagcagcaa	tattgcgctt	cgatgagcct	tggcggtgag	attgatacct	60
	ctgctgcaca	aaaggcaatc	gaccgagctg	gaccagcgca	ttcgtgacac	cgtctccttc	120
	gaacttatcc	gcaatggagt	gtcattcctc	aaggacngcc	tgatcgcaaa	tgggtgctatc	180
	cacgcagcgg	caatcgaaaa	ccttcagccg	gtgaccaata	tctacaacat	cagccttgggt	240
5	atcctgcgtg	atgagccagc	gcagaacaag	gtaaccgcca	gtgccgataa	gttcaaagtt	300
	aaacctgggtg	ttgataccaa	cattgaaacg	ttgatcgaaa	acgcgctgaa	aaacgctgct	360
	gaatgtgcgg	cgctggatgt	cacaaagcaa	atggcagcag	acaagaaagc	gatcgatgaa	420
	ctggcttccct	atgtccgcac	ggccatcatg	atggaatgtt	tccccgggtg	tgttatctgg	480
	cagcagtgcc	gtcgatagta	tgcaattgat	aattattatc	atctgcgggt	cctttccggc	540
10	gatccgcctt	gttacggggc	ggcgacctcg	cgggtttctg	ctatttatga	aaattttccg	600
	gtttaaggcg	tttccgttct	tcttcgtcat	aacttaatgt	ttttatttaa	aataccctct	660
	gaaaagaaaag	gaaacgacag	gtgctgaaaag	cgagcttttt	ggcctctgtc	gtttcctttc	720
	tctgtttttg	tccgtggaat	gaacaatgga	agtcaacaaa	aagcagagct	tatcgatgat	780
	aagcgggtcaa	acatgagaat	tgcgggccgc	ataatacgac	tactatagg	gatcgacgcc	840
15	tactccccgc	gcatgaagcg	gaggagctgg	actccgcctg	cccagagacg	ccccccaacc	900
	cccaaagtgc	ctgacctcag	cctctaccag	ctctggcttg	ggcttggggc	gggtcaaggc	960
	taccacgttc	tcttaacagg	tggctgggct	gtctcttggc	cgcgcgctcat	gtgacagctg	1020
	cctagtctctg	cagtgaggct	accgtggaat	gtctgccttc	gttgccatgg	caacgggatg	1080
	acgtttacaat	ctgggtgtgg	agcttttccct	gtccgtgtca	ggaaatccaa	ataccctaaa	1140
20	ataccctaga	agaggaagta	gctgagccaa	ggctttccctg	gcttctccag	ataaagtttg	1200
	acttagatgg	aaaaaaacaa	aatgataaag	acccgagcca	tctgaaaatt	cctcctaatt	1260
	gcaccactag	gaaatgtgta	tattattgag	ctcgatgtg	ttcttatttt	aaaaagaaaa	1320
	ctttagtcac	gttatttaata	agaatttctc	agcagtggga	gagaaccaat	attaacacca	1380
	agataaaaagt	tggcatgatc	cacattgcag	gaagatccac	gttgggtttt	catgaatgtg	1440
25	aagaccccat	ttattaaagt	cctaagctct	gtttttgcac	actaggaagc	gatggccggg	1500
	atggctgagg	ggctgtaagg	atctttcaat	gtcttacatg	tgtgttttct	gtcctgcacc	1560
	taggacctgc	tgcctagcct	gcagcagagc	cagaggggtt	tcacatgatt	agtctcagac	1620
	acttggggggc	aggttgcatg	tactgcatcg	cctattttcca	tacggagcac	ctactatgtg	1680
	tcaaacacca	tatggtgttc	actcttcaga	acggtgggtg	tcatcatggt	gcatttgctg	1740
30	acggttggat	tgggtgtaga	gagctgagat	atatggacgc	actcttcagc	attctgtcaa	1800
	cgtggctgtg	cattcttgct	cctgagcaag	tggctaaaca	gactcacagg	gtcagcctcc	1860
	agctcagtcg	ctgcatagtc	ttagggaaacc	tctcccagtc	ctccctacct	caactatcca	1920
	agaagccagg	gggcttggcg	gtctcaggag	cctgcttgct	gggggacagg	ttgttgagtt	1980
	ttatctgcag	taggttgccct	aggcatagtg	tcaggactga	tggctgcctt	ggagaacaca	2040
35	tcccttgccc	tctatgcaaa	tctgaccttg	acatggggggc	gctgctcagc	tgggaggatc	2100
	aactgcatac	ctaaagccaa	gcctaaagct	tcttcgtcca	cctgaaactc	ctggaccaag	2160

	gggcttcogg	cacatcctct	caggccagtg	agggagtctg	tgtgagctgc	actttccaat	2220
	ctcagggcgt	gagaggcaga	gggaggtggg	ggcagagcct	tgcagctctt	tcttcccatc	2280
	tggacagcgc	tctggctcag	cagcccatat	gagcacaggc	acatccccac	ccccccccca	2340
	cctttcctgt	cctgcagaat	ttaggctctg	ttcacggggg	gggggggggg	ggggcagtc	2400
5	tatcctctct	taggtagaca	ggactctgca	ggagacactg	ctttgtaaga	tactgcagtt	2460
	taaatttggg	tgttgtagg	ggaaagcgaa	gggcctcttt	gaccattcag	tcaaggtacc	2520
	ttctaactcc	catcgattg	gggggctact	ctagtgcctag	acattgcaga	gagcctcaga	2580
	actgtagtta	ccagtgtggt	aggattgata	cttcagggag	cctgacatgt	gacagttcca	2640
	ttcttcaccc	agtcaccgaa	catttattca	gtacctaccc	cgtaacaggc	accgtagcag	2700
10	gtactgaggg	acggaccact	caaagaactg	acagaccgaa	gccttggaat	ataaacacca	2760
	aagcatcagg	ctctgccaac	agaacactct	ttaacactca	ggccctttta	cactcaggac	2820
	ccccaccccc	accccaagca	gctggcactg	ctatccacat	tttacagaga	ggaaaaacta	2880
	ggcacaggac	gatataagt	gcttgcttaa	gcttgctctg	atggtaaagt	gcagggtcgg	2940
	attgagaccc	agacattcca	actctagggt	ctatttttct	tttttctcgt	tgttcgaatc	3000
15	tgggtcttac	tgggtaaact	caggctagcc	tcacactcat	atccttctcc	catggcttac	3060
	gagtgcctagg	attccaggtg	tgtgctacca	tgtctgactc	cctgtagctt	gtctatacca	3120
	tcctcacaac	ataggaattg	tgatagcagc	acacacaccg	gaaggagctg	gggaaatccc	3180
	acagagggct	ccgcaggatg	acaggcgaat	gcctacacag	aagggtggga	agggaagcag	3240
	agggaacagc	atgggcgtgg	gaccacaagt	ctatttgggg	aagctgccgg	taaccgtata	3300
20	tggctggggg	gaggggagag	gtcatgagat	gaggcaggaa	gagccacagc	aggcagcggg	3360
	tacgggctcc	ttattgccaa	gaggctcgga	tcttcctcct	cttccctcct	ccggggctgc	3420
	ctgttcattt	tccaccactg	cctcccatcc	aggctctgtg	ctcaggacat	caccagctg	3480
	cagaaactgg	gcatcaccca	cgtcctgaat	gctgccgagg	gcaggctcct	catgcacgtc	3540
	aacaccagtg	ctagcttcta	cgaggattct	ggcatcacct	acttgggcat	caaggccaat	3600
25	gatacgagg	agttcaacct	cagtgccttac	tttgaaaggg	ccacagattt	cattgaccag	3660
	gcgctggccc	ataaaaaatg	taaggaacgt	acattccggc	acccatggag	cgtaagccct	3720
	ctgggacctg	cttcctccaa	agaggccccc	acttgaaaaa	ggttccagaa	agatcccaaa	3780
	atatgccacc	aactagggat	taagtgtcct	acatgtgagc	cgatgggggc	cactgcatat	3840
	agctctgtgc	atagacatga	caatggataa	taatatattca	gacagagagc	aggagttagg	3900
30	tagctgtgct	cctttccctt	taattgagt	tgcccatttt	tttattcatg	tatgtgtata	3960
	catgtgtgtg	cacacatgcc	ataggttgat	actgaacacc	gtcttcaatc	gttccccacc	4020
	ccaccttatt	ttttgaggca	gggtctcttc	cctgatcctg	gggctcattg	gtttatctag	4080
	gctgctggcc	agttagctct	ggagtctctg	ttttctctac	ctccctagcc	ctgggactgc	4140
	aggggcatgt	gctgggcccag	gcttttatgt	cgcgttgggg	atctgaactt	aggctccctag	4200
35	gcctgagcac	cgtaaagact	ctgccacatc	cccagcctgt	ttgagcaagt	gaaccattcc	4260
	ccagaattcc	cccagtgggg	ctttcctacc	cttttattgg	ctaggcattc	atgagtggtc	4320



	acctcgccag aggaatgagt ggccacgact ggctcagggc cagcagccta gagatactgg	4380
	gttaagtctt cctgcgcgtc gctccctgca gccgcagaca gaaagtagga ctgaatgaga	4440
	gctggctagt ggtcagacag gacagaaggc tgagaggggtc acagggcaga tgcagcaga	4500
	gcagacaggt tctccctctg tgggggaggg gtggcccact gcagggtgtaa ttggccttct	4560
5	ttgtgctcca tagaggcttc ctgggtacac agcagcttcc ctgtcctggg gattcccaaa	4620
	gagaactccc taccactgga cttacagaag ttctattgac tgggtgtaacg gttcaacagc	4680
	tttggctctt ggtggacggg gcatactgct gtatcagctc aagagctcat tcacgaatga	4740
	acacacacac acacacacac acacacacac acacaagcta attttgatat gccttaacta	4800
	gctcagtgac tgggcatttc tgaacatccc tgaagttagc acacatttcc ctctgggtgt	4860
10	cctggcttaa caccctctaa atctataatt tatctttgct gccctgttac ctctcgagaa	4920
	gcccctaggg ccacttccct tcgcacctac attgctggat ggttctctc ctgcagctct	4980
	taaatctgat cctctctgct ctgagccatg ggaacagccc aataactgag ttagacataa	5040
	aaacgtctct agccaaaact tcagctaaat ttagacaata aatcttactg gttgtggaat	5100
	ccttaagatt ctctcatgac tccctcacat ggcacgagta tgaagcttta ttacaattgt	5160
15	ttattgatca aactaactca taaaaagcca gttgtctttc acctgctcaa ggaaggaaca	5220
	aaattcatcc ttaactgac tgtgcacctt gcacaatcca tacgaataac ttaagagtac	5280
	taagattttg gttgtgagag tcacatgtta cagaatgtac agctttgaca aggtgcatcc	5340
	ttgggatgcc gaagtgcact gctgttccag cccctacctt tctgaggctg ttttgggaagc	5400
	aatgctctgg aagcaacttt aggaggtagg atgctggaac agcgggtcac ttcagcatcc	5460
20	cgatgacgaa tcccgctcaa gctgtacatt ctgtaacaga ctgggaaagc tgcagacttt	5520
	aaggccaggg ccctatgggc cctcttaate cctgtcacac ccaacccgag cccttctcct	5580
	ccagccgttc tgtgcttctc actctggata gatggagaac acggccttgc tagttaaagg	5640
	agtgaggctt cacccttctc acatggcagt ggttgggtcat cctcattcag ggaactctgg	5700
	ggcattctgc ctttacttcc tctttttgga ctacagggaa tatatgctga ctgtttttga	5760
25	ccttgtgtat ggggagactg gatctttggg ctggaatgtt tctgctagt ttttcccat	5820
	cctttggcaa accctatcta tatcttacca ctaggcatag tggccctcgt tctggagcct	5880
	gccttcaggg tggttctcgg ggaccatgtc cctggtttct cccagcata tgggtgttcac	5940
	agtgttcaact gcgggtgggtt gctgaacaaa gcggggattg catcccagag ctccggtgcc	6000
	ttgtgggtac actgctaaga taaaatggat actggcctct ctctgaccac ttgcagagct	6060
30	ctgggtgcctt gtgggtacac tgctaagata aaatggatac tggcctctct ctatccactt	6120
	gcaggactct agggaaacagg aatccattac tgagaaaacc aggggctagg agcagggagg	6180
	tagctgggca gctgaagtgc ttggcgacta accaatgaat accagagttt ggatctctag	6240
	aatactctta aaatctgggt gggcagagtg gcctgcctgt aatcccagaa ctcgagggc	6300
	ggagacaggg aatcatcaga gcaaactggc taaccagaat agcaaaacac tgagctctgg	6360
35	gctctgtgag agatcctgcc ttaacatata agagagagaa taaaacattg aagaagacag	6420
	tagatgccaa ttttaagccc ccacatgcac atggacaagt gtgcgtttga acacacatat	6480

	gcactcatgt gaaccaggca tgcacactcg ggcttatcac acacataatt tgaaagagag	6540
	agtgagagag gagagtgcac attagagttc acaggaaagt gtgagtgagc acacccatgc	6600
	acacagacat gtgtgccagg gagtaggaaa ggagcctggg tttgtgtata agagggagac	6660
	atcatgtgtt tctaaggagg gcgtgtgaag gaggcgttgt gtgggctggg actggagcat	6720
5	ggttgtaact gagcatgttc cctgtgggaa acaggagggg ggccaccctg cagaggggtcc	6780
	cactgtccag cgggatcagt aaaagccctt gctgagaact ttaggtaata gccagagaga	6840
	gaaaggtagg aaagtggggg gactcccatc tctgatgtag gaggatctgg gcaagtagag	6900
	gtgcgtttga ggtagaaaga ggggtgcaga ggagatgttc ttaattctgg gtcagcagtt	6960
	tctttccaaa taatgcctgt gaggaggtgt aggtgggtggc cattcactca ctcagcagag	7020
10	ggatgatgat gcccggtgga tgctggaaat ggccgagcat caaccctggc tctggaagaa	7080
	ctccatcttt cagaaggaga gtggatctgt gtatggccag cggggtcaca ggtgcttggg	7140
	gcccctgggg gactcctagc actgggtgat gtttatcgag tgctcttctg tgccaggcac	7200
	tggcctgggg ctttgtttct gtctctgttt tgtttcgttt tttgagacag actcttgcta	7260
	tgtatccgtg tcaatcttgg aatctcactg catagcccag gctgcggaga gaggggaggg	7320
15	caataggcct tgtaagcaag ccacacttca gagactagac tccaccctgc gaatgatgac	7380
	aggtcagagc tgagttccgg aagatttttt tccagctgc cagggtggagt gtggagtggc	7440
	agctagcggc aagggtagag ggcgagctcc ctgtgcagga gaaatgcaag caagagatgg	7500
	caagccagtg agttaagcat tctgtgtggg gagcaggtgg atgaagagag aggctgggct	7560
	ttcgccctctg gggggggggg gaggggtggg gatgaggtga gagggggca gctccctgca	7620
20	gtgtgatgag atttttcctg acagtgacct ttggcctctc cctccccac ttcccttctt	7680
	tcctttcttc ccaccattgc ttcccttctc cttgagaaat tctgagtttc cacttcactg	7740
	gtgatgcaga cggaaacaga agccgtgtgt gtgtgtgtgt gtgtgtgtgt gtgtgtgtgt	7800
	gtgtgtgtgt ttgtgtgtat gtgtgtgtgt gtgtttgtgt gtatgtgtgt cagtgggaat	7860
	ggctcatagt ctgcaggaag gtgggcagga aggaataagc tgtaggctga ggcagtgtgg	7920
25	gatgcagggg gagaggagag gagggatacc agagaaggaa attaaggag ctacaaggag	7980
	gcattgttgg ggtgtgtgtg tgtgtgtgtt gtttatattt gtattggaaa tacattcttt	8040
	taaaaaatac ttatccattt atttatattt atgtgcacgt gtgtgtgcct gcatgagttc	8100
	atgtgtgcca cgtgtgtgcg ggaacccttg gaggccacaa gggggcatct gatcccttgg	8160
	aactggagtt ggaggaggtt gtgagtcctc tgacatgttt gctgggaact gaaccccggt	8220
30	cctatgcaag agcaggaagt gcagttatct gctgagccat ctctccagtc ctgaaatcca	8280
	ttctcttaaa atacacgtgg cagagacatg atgggattta cgtatggatt taatgtggcg	8340
	gtcattaagt tccggcacag gcaagcacct gtaaagccat caccacaacc gcaacagtga	8400
	atgtgaccat ccccccatg ttcttcatgt cccctgtccc ctccatctc cattctcaag	8460
	cacctcttgc tctgcctctg tcgctggaga acagtgtgca tctgcacact cttatgtcag	8520
35	tgaagtcaca cagcctgcac ccttccttgg tctgagtatt tgggttctga ctctgctatc	8580
	acacactact gtactgcatt ctctcgctct ctttttttaa acatattttt atttgtttgt	8640

	gtgtatgcac atgtgccaca tgtgtacaga tactatggag gccagaagag gccatggccg	8700
	tccctggagc tggagttaca ggcagcgtgt gagctgcctg gtgtgggtgc tgggaaccaa	8760
	acttgaatct aaagcaagca cttttaactg ctgaggcagc tctcagtacc cttcttcatt	8820
	tctccgcctg ggttccattg tatggacaca tgtagctaga atatcttgct tatctaatta	8880
5	tgtacattgt tttgtgctaa gagagagtaa tgctctatag cctgagctgg cctcaacctt	8940
	gccatcctcc tgcctcagcc tccctcctct gagtgctagg atgacaggcg agtggttaact	9000
	tacatgggtt catgttttgt tcaagactga aggataacat tcatacagag aagggtctggg	9060
	tcacaaagtg tgcagttcac tgaatggcac aaccctgat caagaaacaa aactcagggg	9120
	ctggagagat ggcactgact gctcttccag aggtccggag ttcaattccc agcaaccaca	9180
10	tgggtggctca cagccatcta taacgagatc tgacgcctc ttctgggtgtg tctgaagaca	9240
	gctacagtgt actcacataa aataaataaa tctttaaaac acacacacac acacaattac	9300
	caccccagaa agccactcc atgttccctc ccacgtctct gcttacagta ctcccaggtt	9360
	accactgttc aggtttctaa caacctgggt tacttggggc tcttttctgc tctgtggagc	9420
	cacacatttg tgtgcctcat acacgttctt tctagtaagt tgcataattac tctgcgtttt	9480
15	tacatgtatt tattttattgt agttgtgtgt gcgtgtgggc ccatgcatgg cacagtgtgt	9540
	ggggatgtca gagtatttg aacaggggac agttcttttc ttcaatcatg tgggttccag	9600
	aggttgaact caggtcatca tgtgtggcag caaatgcctt taccactga gacatctcca	9660
	tattcttttt ttttccctg aggtgggggc ttgttccata gcccactg gctttgact	9720
	tgcagttcaa agtgactccc tgtctccacc tcttagagta ttggaattac gatgtgtact	9780
20	accacacctg actggatcat taattctttg atgggggcgg ggaagcgcac atgctgcagg	9840
	tgaagggatg actggactgg acatgagcgt ggaagccaga gaacagcttc agtctaattgc	9900
	tctcccaact gagctatttc ggtttgccag agaacaactt acagaaagtt ctcatgcca	9960
	tgtggattcg gggttggagt tcaactcatc agcttgacat tggctcctct acccactgag	10020
	ccttctcact actctctacc tagatcatta attctttttt aaaaagactt attagggggc	10080
25	tggagagatg gctcagccgt taagagcacc gaatgccctt ccagaggctc tgagttcaat	10140
	tcccagcatg ccattgctgg gcagtagggg gcgcaggtgt tcaacgtgag tagctgttgc	10200
	cagttttccg cgggtggagaa cctcttgaca cctgctgtc cctggtcatt ctgggtgggt	10260
	gcatggtgat atgcttggtg tatggaagac ttgtactgtt acagtgaagt tgggcttcca	10320
	cagttaccac gtctcccctg tttcttgag gccgggtgct tgtccattgc cgcgagggt	10380
30	acagccgtc cccaacgcta gttatgcct acctcatgat gcggcagaag atggacgtca	10440
	agtctgctct gagtactgtg aggcagaatc gtgagatcgg cccaacgat ggcttccctg	10500
	cccaactctg ccagctcaat gacagactag ccaaggaggg caaggtgaaa ctctaggggtg	10560
	cccacagcct cttttgcaga ggtctgactg ggagggccct ggcagccatg tttaggaaac	10620
	acagtatacc cactcccctg accaccagac acgtgccac atctgtccca ctctggtcct	10680
35	cggggggcac tccaccctta gggagcacat gaagaagctc cctaagaagt tctgctcctt	10740
	agccatcctt tccgttaatt tatgtctctc cctgagggtga gggtcagggt tatgtccctg	10800

	tctgtggcat agatacatct cagtgaacca ggggtggagg gctatcaggg tgcattggccc	10660
	gggacacggg cactcttcat gacccctccc ccacctgggt tcttctgtg tgggtccagaa	10920
	ccacgagcct ggtaaaggaa ctatgcaaac acaggccctg acctcccat gtctgttctt	10980
	ggctcctcaca gcccagacacg cctgtgtgag gcagacgaat gacattaagt tctgaagcag	11040
5	agtgagagata gattagtga tagatttcca aaaagaagga aaaaaaaggc tgcattttaa	11100
	aattatttcc ttagaattaa agatactaca taggggcccct tgggtaagca aatccatttt	11160
	tcccagaggc tatcttgatt ctttggaatg tttaaagtgt gccttgccag agagcttacg	11220
	atctatatct gctgcttcag agccttccct gaggatggct ctgttccctt gcttgtaga	11280
	agagcgatgc cttgggcagg gtttccccct tttcagaata cagggtgtaa agtccagcct	11340
10	attacaaaca aacaaacaaa caaacaacaa aaggacctcc atttggagaa ttgcaaggat	11400
	tttatcctga attatagtgt tgggtgagttc aagtcacac gccaaagtgt tgcctcctg	11460
	gttgctattc taagaataat taggaggagg aacctagcca attgcagctc atgtccgtgg	11520
	gtgtgtgcac ggggtgcata gttaggaagg gtgcctgtcc ccttggggac agaaggaaaa	11580
	tgaaggccc ctctgtcac cctggccatt tacgggaggc tctgtgtgtt ccacgggtgtc	11640
15	tgtgcaggat cctgaaactg actcgttga cagaaacgag acctggcggc accatgagaa	11700
	tggagagaga gagagcaaag aaagaaacag cctttaaag aactttctaa ggggtgtttt	11760
	tgaacctcgc tggacctgt atgtgtgcac atttgccaga gattgaacat aatcctcttg	11820
	ggacttcacg ttctcattat ttgtatgtct ccggggtcac gcagagccgt cagccaccac	11880
	cccagcaccc ggcacatagg cgtctcataa aagcccattt tatgagaacc agagctgttt	11940
20	gagtacccc tgtatagaga gagttgtgt cgtggggcac ccggatccca gcagcctggt	12000
	tgcctgcctg taggatgtct tacaggagtt tgcagagaaa ccttcccttg agggaaagaa	12060
	atatcaggga tttttgttga atatttcaaa ttcagcttta agtgaagac tcagcagtgt	12120
	tcattggttaa ggtaaggaa atgccttttc cagagctgtt gcaagaggca ggagaagcag	12180
	acctgtctta ggatgtcact cccagggtta agacctctga tcacagcagg agcagagctg	12240
25	tgcagcctgg atggctattg tcccctattc tgtgtgacca cagcaacct ggtcacatag	12300
	ggctggctat cctttttttt tttttttttt ttttttttg gcccagaatg aagtgacct	12360
	agccaagttg tgtacctcag tctttagttt ccaagcggct ctcttgctca atacaatgtg	12420
	catttcaaaa taacactgta gagttgacag aactggttca tgtgttatga gagaggaaaa	12480
	gagaggaaa aacaaaacaa aacaaaacac caciaaccaa aaacatctgg gctagccagg	12540
30	catgattgca atgtctacag gccagttca tgagaggcag agacaggaag accgccgaaa	12600
	ggtcaaggat agcatggctt acgtatcgag actccagcca gggctacggt cccaagatcc	12660
	taggttttgg attttgggct ttggtttttg agacagggtt tctctgtgta gccctggctg	12720
	tcctggaact cgctctgtag accaggctgg cctcaaaact agagatctgc ctgactctgc	12780
	ctttgagggc tgggacgaat gccaccactg cccaactaag attccattaa aaaaaaaaaa	12840
35	agttcaagat aattaagagt tgccagctcg ttaaagctaa gtagaagcag tctcaggcct	12900
	gctgcttgag gctgttcttg gcttggacct gaaatctgcc cccaacagtg tccaagtgca	12960

	catgactttg agccatctcc agagaaggaa gtgaaaattg tggctcccca gtcgattggg	13020
	acacagtctc tctttgtcta ggtaacacat ggtgacacat agcattgaac tctccactct	13080
	gaggggtgggt ttcctccccc ctgcctcttc tgggttggtc accccatagg acagccacag	13140
	gacagtcaact agcacctact ggaaacctct ttgtgggaac atgaagaaaag agcctttggg	13200
5	agattcctgg ctttccatta gggctgaaag tacaacgggt cttgggtggc tttgcctcgt	13260
	gtttataaaa ctagctacta tctttcaggt aaaataccga tgtgtggaa aagccaaccc	13320
	cgtggctgcc cgtgagtagg ggggtggggtt ggggaatcctg gatagtgttc tatccatgga	13380
	aagtggtgga ataggaatta aggggtgttc cccccccccc aacctcttcc tcagacccag	13440
	ccactttcta tgacttataa acatccaggt aaaaattaca aacataaaaa tgggttctct	13500
10	tctcaatctt ctaaagtctg cctgcctttt ccaggggtag gtctgtttct ttgctgttct	13560
	attgtcttga gagcacagac taacacttac caaatgaggg aactcttggc ccataactaag	13620
	gctcttctgg gctccagcac tcttaagtta ttttaagaat tctcacttgg cctttagcac	13680
	acccgccacc cccaagtggg tgtggataat gccatggcca gcagggggca ctgttgaggc	13740
	gggtgccttt ccaccttaag ttgcttatag tatttaagat gctaaatgtt ttaatcaaga	13800
15	gaagcactga tcttataata cgaggataag agattttctc acaggaaatt gtctttttca	13860
	taattctttt acaggctttg tcttgatcgt agcatagaga gaatagctgg atatttaact	13920
	tgtattccat tttcctctgc cagcgttagg ttaactccgt aaaaagtgat tcagtggacc	13980
	gaagaggctc agagggcagg ggatggtagg gtgaggcaga gcactgtcac ctgccaggca	14040
	tgggaggtcc tgccatccgg gaggaagg aaagtttagc ctctagtcta ccaccagtgt	14100
20	taacgcactc taaagtgtga accaaaaata atgtcttaca ttacaaagac gtctgttttg	14160
	tgtttccttt tgtgtgtttg ggctttttat gtgtgcttta taactgctgt ggtggtgctg	14220
	ttgttagttt tgaggtagga tctcaggctg gccttgaact tctgatcgcc tgcctctgcc	14280
	cctgccccctg cccctgtccc tgctccaag tgctaggact aaaagcacat gccaccacac	14340
	cagtacagca tttttctaac atttaaaaat aatcacctag gggctggaga gagggttcca	14400
25	gctaagagtg cacactgtct ttgggttagga cctgagttta gtcccagaa cctatactgg	14460
	gtggctccag gtccagagga tccaggacct ctggcctcca tgggcatctg ctcttagcac	14520
	ataccacat acagatacac acataaaaat aaaatgaagc ctttaaaaac ctctaaaac	14580
	ctagcccttg gaggtacgac tctggaaagc tggcacttg tgtaagtcca tctcatggtg	14640
	ttctggctaa cgtaagactt acagagacag aaaagaactc aggggtgtgct gggggttggg	14700
30	atggaggaag agggatgagt agggggagca cggggaactt gggcagtga aattctttgc	14760
	aggacactag agggaggata ataccagtca ttgcacccac tactggacaa ctccaggga	14820
	ttatgctggg tgaaaagaga agggccccagg tattggctgc attggctgca tttgcgtaac	14880
	atttttttaa attgaaaaga aaaagatgta aatcaagggt agatgagtgg ttgctgtgag	14940
	ctgagagctg ggggtgagtga gacatgtgga caactccatc aaaaagcgac agaaagaacg	15000
35	ggctgtggtg acagctacct ctaatctcca cctccgggag gtgatcaagg ttagccctca	15060
	gctagcctgt ggtgcatgag accctgtttc aaaaacttta ataaagaaat aatgaaaaaa	15120

	gacatcaggg cagatccttg gggccaaagg cggacaggcg agtctcgtgg taaggctgtg	15180
	tagaagcggg tgcattgagc cgtgcccgag gcatcatgag agagccctag gtaagtaagg	15240
	atggatgtga gtgtgtcggc gtcggcgac tgcacgtcct ggctgtggtg ctggactggc	15300
	atcttttggtg agctgtggag gggaaatggg tagggagatc ataaaatccc tccgaattat	15360
5	ttcaagaact gtctattaca attatctcaa aatattaaaa aaaaagaaga attaaaaaac	15420
	aaaaaaccta tccagggtgtg gtgggtgtgca cctatagcca cgggcacttg gaaagctgga	15480
	gcaagaggat ggcgagtttg aaggatcttg gggctgtaca gcaagaccgt cgtcccaaaa	15540
	ccaaacaaaa cagcaaaccc attatgtcac acaagagtgt ttatagttag cggcctcgct	15600
	gagagcatgg ggtgggggtg ggggtggggg acagaaatat ctaaactgca gtcaataggg	15660
10	atccactgag accctggggc ttgactgcag cttaaccttg ggaaatgata agggtttgt	15720
	gttgagttaa agcatcgatt actgacttaa cctcaaatga agaaaaagaa aaaaagaaaa	15780
	caacaaaagc caaaccaagg ggctgggtgag atggctcagt gggtaagagc acccgactgc	15840
	tcttccgaag gtccagagtt caaatcccag caaccacatg gtggctcaca accatctgta	15900
	acgagatatg atgccctctt ctgggtgtgtc tgaagacagc tacagtgtac ttacatataa	15960
15	taaataaatc ttaaaaaaaa aaaaaaaaaa aaaagccaaa ccgagcaaac caggcccca	16020
	aacagaaggc aggcacgacg gcaggcacca cgagccatcc tgtgaaaagg cagggtacc	16080
	catgggccga ggagggtcca gagagatagg ctggtaagct cagtttctct gtataccctt	16140
	tttcttgttg acactacttc aattacagat aaaataacaa ataaacaaaa tctagagcct	16200
	ggccactctc tgcctcgttg atttttcttg ttacgtccag cagggtggcg aagtgttcca	16260
20	aggacagatc gcatcattaa ggtggccagc ataatctccc atcagcaggt ggtgctgtga	16320
	gaaccattat ggtgctcaca gaatccccggg cccaggagct gccctctccc aagtctggag	16380
	caataggaaa gctttcttggc ccagacaggg ttaacagtcc acattccaga gcaggggaaa	16440
	aggagactgg aggtcacaga caaaagggcc agcttctaac aacttcacag ctctggtagg	16500
	agagatagat ccccccaaac aatggccaca gctgggtttg tctgccccga aggaaactga	16560
25	cttaggaagc aggtatcaga gtcccccttc tgaggggact tctgtctgcc ttgtaaagct	16620
	gtcagagcag ctgcattgat gtgtgggtga cagaagatga aaaggaggac ccaggcagat	16680
	cgccacagat ggaccggcca cttaacagtc gaggcagggt gcagagcctt gcagaagctc	16740
	tgcagggtga cgacactgat tcattaccca gttagcatat cacagcgggc taggcggacc	16800
	acagcctcct tcccagttct cctccagggc tggggaggtc tccaaccttc tgtctcagt	16860
30	cagcttccgc cagccccctc tctttttgca cctcagggtg gaacctccc tctctcctt	16920
	ctccctgtgg catggccctc ctgctactgc aggtcagca ttggatttct ttgtgcttag	16980
	atagacctga gatggcttct tgatttatat atatatatcc atcccttga tcttacatct	17040
	aggacctcga gctgtttgtg ataccataag aggtcgggga gatgatatgg taagagtgtc	17100
	tgtgtacaa gcatgaagac atgagttcga atccccagca accatgtgga aaaataacct	17160
35	tctaacctca gagttgaggg gaaaggcagg tggattctgg gggcttactg gccagctagc	17220
	cagcctaacc taaatgtctc agtcagagat cctgtctcag ggaataactt gggagaatga	17280

	ctgagaaaga cacttccctca ggtctcccat gcacccacac agacacacgg ggggggggta	17340
	atgtaataag ctaagaaata atgagggaaa tgattttttg ctaagaaatg aaattctgtg	17400
	ttggccgcaa gaagcctggc caggggaagga actgcctttg gcacaccage ctataagtca	17460
	ccatgagttc cctggctaag aatcacatgt aatggagccc aggtccctct tgccctgggg	17520
5	ttgcctctcc cactggtttt gaagagaaat tcaagagaga tctccttggg cagaattgta	17580
	gggtgctgagc aatgtggagc tgggggtcaat gggattccct taaaggcatc cttcccaggg	17640
	ctgggtcata cttcaatagt aggggtgcttg cacagcaagc gtgagaccct aggttagagt	17700
	ccccagaatc tgccccaac cccccaaaaa ggcatecttc tgccctctggg tgggtggggg	17760
	gagcaaacac ctttaactaa gaccattagc tggcaggggt aacaaatgac cttggctaga	17820
10	ggaatttggg caagctggat tccgccttct gtagaagccc cacttgtttc ctttgtaag	17880
	ctggcccaca gtttggtttg agaatgcctg aggggcccag ggagccagac aattaaaagc	17940
	caagctcatt ttgatattctg aaaaccacag cctgactgcc ctgcccgtgg gaggtactgg	18000
	gagagctggc tgtgtccctg cctcaccaac gccccccccc ccaacacaca ctccctcggt	18060
	cacctgggag gtgccagcag caatttgga gtttactgag cttgagaagt cttgggaggg	18120
15	ctgacgctaa gcacaccctt tctccacccc cccccacccc acccccgtga ggaggagggg	18180
	gaggaaacat gggaccagcc ctgctccagc ccgtccttat tggctggcat gaggcagagg	18240
	gggctttaaa aaggcaaccg tatctaggtt ggacactgga gcctgtgcta ccgagtcccc	18300
	tctccacctt ggcagcatgc agccctcact agccccgtgc ctcatctgcc tacttgtgca	18360
	cgctgccttc tgtgtgtgtg agggccaggg gtggcaagcc ttcaggaatg atgccacaga	18420
20	ggtcatccca gggcttggag agtaccctga gcctcctcct gagaacaacc agaccatgaa	18480
	ccgggcgag aatggaggca gacctccca ccatccctat gacgcaaag gtacgggatg	18540
	aagaagcaca ttagtggggg ggggggtcct gggaggtgac tgggggtggt ttagcatctt	18600
	cttcagaggt ttgtgtgggt ggctagcctc tgctacatca gggcagggac acatttgctt	18660
	ggaagaatac tagcacagca ttagaacctg gagggcagca ttgggggggct ggtagagagc	18720
25	acccaaggca ggggtggaggc tgaggtcagc cgaagctggc attaacacgg gcatgggctt	18780
	gtatgatggt ccagagaatc tcttcctaag gatgaggaca caggtcagat ctagtgtgtg	18840
	accagtgggg aagtgatatg gtgaggctgg atgccagatg ccatccatgg ctgtactata	18900
	tcccacatga ccaccacatg aggtaaagaa gggcccagct tgaagatgga gaaaccgaga	18960
	ggctcctgag ataaagtcac ctgggagtaa gaagagctga gactggaagc tggtttgatc	19020
30	cagatgcaag gcaaccctag attgggtttg ggtgggaacc tgaagccagg aggaatccct	19080
	ttagttcccc cttgcccagg gtctgctcaa tgagcccaga gggttagcat taaaagaaca	19140
	gggtttgtag gtggcatgtg acatgagggg cagctgagtg aaatgtcccc tgtatgagca	19200
	caggtggcac cacttgccct gagcttgac cctgacccca gctttgcctc attcctgagg	19260
	acagcagaaa ctgtggaggc agagccagca cagagagatg cctgggggtgg ggggtggggg	19320
35	atcacgcacg gaactagcag caatgaatgg ggtgggggtgg cagctggagg gacactccag	19380
	agaaatgacc ttgctgggtca ccatttgtgt gggaggagag ctcattttcc agcttgccac	19440

	cacatgctgt cccctcctgtc tccatagccag taagggatgt ggaggaaaagg gccaccccaa	19500
	aggagcatgc aatgcagtcg cgtttttgca gaggaagtgc ttgacctaaag ggcactattc	19560
	ttggaaagcc ccaaaactag tccctccctg ggcaaacagg cctccccccac ataccacctc	19620
	tgcaggggtg agtaaattaa gccagccaca gaaggggtggc aaggcctaca cctccccctt	19680
5	gttgtgcccc cccccccccc gtgaagggtgc atcctggcct ctgccccctt ggcttttggtg	19740
	ctgggatttt ttttttccctt ttatgtcata ttgatcctga caccatggaa cttttggagg	19800
	tagacaggac ccacacatgg attagttaaa agcctcccat ccatctaagc tcatggtagg	19860
	agatagagca tgtccaagag aggagggcag gcatcagacc tagaagatat ggctgggcat	19920
	ccaacccaat ctccttcccc ggagaacaga ctctaagtca gatccagcca cccttgagta	19980
10	accagctcaa ggtacacaga acaagagagt ctgggtatata gcagggtgcta aacaaatgct	20040
	tgtggttagca aaagctatag gttttgggtc agaactccga cccaagtgcg gagtgaagag	20100
	cgaaaggccc tctactcgcc accgccccgc ccccacctgg ggtcctataa cagatcactt	20160
	tcacccctgc gggagccaga gagccctggc atcctaggta gcccccccg cccccccccc	20220
	gcaagcagcc cagccctgcc tttggggcaa gttcttttct cagcctggac ctgtgataat	20280
15	gaggggggtg gacgcgcgc ctttgggtgc tttcaagtct aatgaattct tatccctacc	20340
	acctgcccc ctacccccgt cctccacagc agctgtcctg atttattacc ttcaattaac	20400
	ctccactcct ttctccatct cctgggatac cgccccgtc ccagtggctg gtaaaggagc	20460
	ttaggaagga ccagagccag gtgtggctag aggtaccag gcagggtgg ggatgaggag	20520
	ctaaactgga agagtgtttg gttagtaggc acaaagcctt ggggtgggac cctagtaccg	20580
20	gagaagtgga gatgggcgt gagaagttca agaccatcca tccttaacta cacagccagt	20640
	ttgaggccag cctgggctac ataaaaaccc aatctcaaaa gctgccatt ctgattctgt	20700
	gccacgtagt gcccgatgta atagtggatg aagtcgttga atcctggggc aacctatatt	20760
	acagatgtgg ggaaaagcaa ctttaagtac cctgcccaca gatcaciaag aaagtaagtg	20820
	acagagctcc agtgtttcat ccttgggttc caaggacagg gagagagaag ccagggtggg	20880
25	atctcactgc tccccggtgc ctcttcccta taatccatac agattcgaaa gcgcagggca	20940
	ggtttggaaa aagagagaag ggtggaagga gcagaccagt ctggcctagg ctgcagcccc	21000
	tcacgcattc ctctctccgc agatgtgtcc gagtacagct gccgcgagct gcactacacc	21060
	cgcttccctga cagacggccc atgccgcagc gccaaagccg tcaccgagtt ggtgtgctcc	21120
	ggccagtgcg gccccgcgcg gctgctgccc aacgccatcg ggcgcgtaga gtggtggcgc	21180
30	ccgaacggac cggatttccg ctgcatcccg gatcgctacc gcgcgcagcg ggtgcagctg	21240
	ctgtgccccg ggggcgcggc gccgcgctcg cgcaagggtg gtctgggtggc ctctgcaag	21300
	tgcaagcgc tcacccgctt ccacaaccag tcggagctca aggacttcgg gccggagacc	21360
	gcgcggccgc agaagggtcg caagccgcgg cccggcgccc ggggagccaa agccaaccag	21420
	gcggagctgg agaacgccta ctagagcgag cccgcgccta tgcagcccc gcgcgatccg	21480
35	attcgttttc agtgtaaagc ctgcagccca ggccaggggt gccaaacttt ccagaccgtg	21540
	tggagtcccc agcccagtag agaccgcagg tccttctgcc cgctgcgggg gatggggagg	21600



	gggtgggggt cccgcggggc aggagaggaa gctttagtcc cagactctgc ctagccccgg	21650
	gtgggatggg ggtcttttca ccttcgcccg acctatacag gacaaggcag tgtttccacc	21720
	ttaaagggaa gggagtgtgg aacgaaagac ctgggactgg ttatggacgt acagtaagat	21780
	ctactccttc caccctaatg taaagcctgc gtgggctaga tagggtttct gacctgacc	21840
5	tggccactga gtgtgatgtt gggctacgtg gtctcttttt ggtacgggtct tctttgtaaa	21900
	atagggaccg gaactctgct gagattccaa ggattgggggt accccgtgta gactggtgag	21960
	agagaggaga acaggggagg ggttagggga gagattgtgg tgggcaaccg cctagaagaa	22020
	gctgtrttgtt ggctcccagc ctgcgcgcct cagaggtttg gcttccccca ctcttccttc	22080
	tcaaactctgc cttcaaactc atatctggga tagggaaggc cagggtccga gagatggtgg	22140
10	aagggccaga aatcacactc ctggcccccc gaagagcagt gtcccgcccc caactgcctt	22200
	gtcatattgt aaagggattt tctacacaac agtttaaggt cgttggagga aactgggctt	22260
	gccagtcacc tcccctcctt gtcccttgcc aggacaccac ctctgcctg ccaccacagg	22320
	acacatttct gtctagaaac agagcgtcgt cgtgctgtcc tctgagacag catatcttac	22380
	attaaaaaga ataatacggg gggggggggc ggagggcgca agtggtatac atatgctgag	22440
15	aagctgtcag gcgccacagc accaccaca atctttttgt aaatcatttc cagacacctc	22500
	ttactttctg tgtagatttt aattgttaaa aggggaggag agagagcgtt tgtaacagaa	22560
	gcacatggag gggggggtag ggggggtggg gctggtagt ttggcgaact ttccatgtga	22620
	gactcatcca caaagactga aagccgcgtt ttttttttta agagttcagt gacatattta	22680
	ttttctcatt taagtatttt atgccaatat tttttctctg tagagaaagg cagtgttaat	22740
20	atcgctttgt gaagcacaaag tgtgtgtggt tttttgtttt ttgttttttc cccgaccaga	22800
	ggcattgtta ataaagacaa tgaatctcga gcaggaggct gtggtcttgt tttgtcaacc	22860
	acacacaatg tctcgccact gtcattctac tcccttcctt tggtcacaag acccaaacct	22920
	tgacaacacc tccgactgct ctctggtagc ccttgtggca atacgtgttt cctttgaaaa	22980
	gtcacattca tcccttcctt tgcaaacctg gctctcattc cccagctggg tcatcgtcat	23040
25	acctcaccc cagcttcctt ttagctgacc actctccaca ctgtcttcca aaagtgcacg	23100
	tttcaccgag ccagttccct ggtccaggct atcccattgc tccctccttg tccagacctt	23160
	tctcccacaa agatgttcat ctcccactcc atcaagcccc agtggccctg cggctatccc	23220
	tgtctcttca gttagctgaa tctacttget gacaccacat gaattccttc cctgtctta	23280
	aggttcatgg aactcttgcc tgccccgaa ccttcagga ctgtcccagc gtctgatgtg	23340
30	tctctctctt tgtaaagccc caccctacta tttagattccc aattctagat ctcccttgt	23400
	tcattccttc acgggatagt gtctcatctg gccaaagtcct gcttgatatt gggataaatg	23460
	caaagccaag tacaattgag gaccagttca tcattggggc aagctttttc aaaatgtgaa	23520
	ttttacacct atagaagtgt aaaagccttc caaagcagag gcaatgcctg gctcttcctt	23580
	caacatcagg gctcctgctt tatgggtctg gtggggtagt acattcataa acccaacact	23640
35	aggggtgtga aagcaagatg attgggagtt cgaggccaat cttggctatg aggcctgtc	23700
	tcaacctctc ctccctccct ccagggtttt gttttgtttt gtttttttga tttgaaactg	23760

	caacacttta aatccagtc aatgcacatct tgcgtgaggg gaactctatc cetaatatataa	23820
	gcttccatct tgatttgtgt atgtgcacac tgggggttga acctgggcct ttgtacctgc	23880
	cgggcaagct ctctactgct ctaaaccag cctcactgg ctttctgttt caactcccaa	23940
	tgaattcccc taaatgaatt atcaatatca tgtctttgaa aaataccatt gagtgtctgt	24000
5	gggtgccccg tgggtccaga tcccaggaag gacttttcag ggaatccagg catcctgaag	24060
	aatgtcttag agcaggaggc catggagacc ttggccagcc ccacaaggca gtgtggtgca	24120
	gaggggtgagg atggaggcag gcttgcaatt gaagctgaga cagggtactc aggattaaaa	24180
	agcttcccc aaacaattc caagatcagt tcttggtact tgcacctgtt cagctatgca	24240
	gagcccagtg ggcataagtg aagacaccgg ttgtactgtc atgtactaac tgtgcttcag	24300
10	agccggcaga gacaaataat gttatggtga ccccagggga cagtgattcc agaaggaaca	24360
	cagaagagag tgctgctaga ggctgcctga aggagaaggg gtcccagact ctctaagcaa	24420
	agactccact cacataaaga cacaggctga gcagagctgg ccgtggatgc agggagccca	24480
	tccaccatcc tttagcatgc ccttgtattc ccatcacatg ccagggatga ggggcatcag	24540
	agagtccaag tgatgccccaa acccaaacac acctaggact tgctttctgg gacagacaga	24600
15	tgcaggagag actaggttgg gctgtgatcc cattaccaca aagagggaaa aaacaaaaaa	24660
	caaacaaaca aacaaaaaaa aacaaaacaa aacaaaaaaa aaccaaggt ccaaattgta	24720
	ggtcagggtta gagtattatt atggaaagtt atattctacc tccatgggggt ctacaaggct	24780
	ggcgcccatc agaaagaaca aacaacaggc tgatctggga ggggtggtac tctatggcag	24840
	ggagcacgtg tgcttggggg acagccagac acggggcttg tattaatcac agggcttgta	24900
20	ttaataggct gagagtcaag cagacagaga gacagaagga aacacacaca cacacacaca	24960
	cacacacaca cacacacaca catgcacaca cactcactt ctactcgaa gagccccctac	25020
	ttacattcta agaacaaacc attctctctc ataaaggaga caaagttgca gaaacccaaa	25080
	agagccacag ggtccccact ctctttgaaa tgacttggaac ttgttgagg gaagacagag	25140
	gggtctgcag aggtctcctg ggtgaccag agccacagac actgaaatct ggtgctgaga	25200
25	cctgtataaa cctcttcca caggttcctt gaaaggagcc cacattcccc aacctgtct	25260
	cctgaccact gaggatgaga gcacttgggc ctccccatt ctgggagtgc acctggttt	25320
	ccccatctga gggcacatga ggtctcaggt cttgggaaag tccacaagt attgaaagtg	25380
	ttcttggttt gtttgtgatt taatttaggt gtatgagtgc ttttgcttga atatatgcct	25440
	gtgtagcatt tacaagcctg gtgcctgagg agatcagaag atggcatcag atacctgga	25500
30	actggacttg cagacagtta tgagccactg tgtgggtgct aggaacagaa cctggatcct	25560
	ccggaagagc agacagccag cgctcttagc cactaagcca tctactgaggt tctttctgtg	25620
	gctaaagaga caggagacaa aggagagttt ctttttagtca ataggaccat gaatgttct	25680
	cgtaacgtga gactagggca ggggtgatcc ccagtgacac cgatggccct gtgtagttat	25740
	tagcagctct agtcttattc cttaataagt ccagtttgg ggcaggagat atgtattccc	25800
35	tgctttgaag tggctgaggt ccagttatct acttccaagt acttgtttct ctttctggag	25860
	ttggggaagc tccctgcctg cctgtaaatg tgtccattct tcaaccttag acaagatcac	25920

	tttccctgag cagtcaggcc agtccaaagc ccttcaattt agctttcata aggaacaccc	25980
	cttttggttg gtggaggtag cacttgccct gaatccagc attaagaagg cagagacagt	26040
	cggatctctg tgagttcaca gccagcctgg tctacggagt gagttccaag acagccaggc	26100
	ctacacagag aaacctgtc tcgaaaaaaa caaaaacaaa agaaataaag aaaaagaaaa	26160
5	caaaaacgaa caaacagaaa aacaagccag agtggttgtc cccgtatttt attaatcata	26220
	tttttgctcc ttgcccattt tagactaaaa gactcgggaa agcagggtctc tctctgttcc	26280
	tcacccggac acaccagaa ccagatgtat ggaagatggc taatgtgctg cagttgcaca	26340
	tctggggctg ggtggattgg ttagatggca tgggctgggt gtggttacga tgactgcagg	26400
	agcaaggagt atgtggtgca tagcaaacga ggaagtttgc acagaacaac actgtgtgta	26460
10	ctgatgtgca ggtatgggca catgcaagca gaagccaagg gacagcctta gggtagtgtt	26520
	tccacagacc cctccccctt tttaacatgg gcattcttca ttggcctgga gcttgccaac	26580
	tgggctgggc tggctagctt gtaggttcca gggatctgca tatctctgcc tccctagtgc	26640
	tgggattaca gtcatatatg agcacacctg gcttttttat gtgggttctg ggctttgaac	26700
	ccagatctga gtgcttgcaa ggcaatcggg tgaatgactg cttcatctcc ccagaccctg	26760
15	ggattctact ttctattaaa gtatttctat taaatcaatg agccccctgcc cctgcactca	26820
	gcagttctta ggctgctga gagtcaagtg gggagtgaga gcaagcctcg agaccccatc	26880
	agcgaagcag aggacaaaga aatgaaaact tgggattcga ggctcgggat atggagatac	26940
	agaaagggtc aggggaaggaa atgaaccaga tgaatagagg caggaagggt agggccccgc	27000
	atacatggaa cctgggtgtac atgttatctg catgggggtt gcattgcaat ggctcttcag	27060
20	caggttcacc aactgggaa acagaagcca aaaagaagag taggtggtgt tggagtcaga	27120
	tactgtcagt catgcctgaa gaaatggaag caattaacga tgcgccgcaa ttaggatatt	27180
	agctccctga agaaaggcaa gaagctgggc tgtgggcact gaaggagct ttgaatgatg	27240
	tcacattctc tgtatgccta gcagggcagt attggagact gagacttgac ttgtgtgtcc	27300
	atatgattcc tctttttcct acagtcactt ggggctcctg agcttcgtcc ttgtccaaga	27360
25	acctggagct ggcagtgggc agctgcagtg atagatgtct gcaagaaaga tctgaaaaga	27420
	gggaggaaga tgaaggacc agaggaccac cgacctctgc tgccctgacaa agctgcagga	27480
	ccagtccttc ctacagatgg gagacagagg cgagagatga atggtcaggg gaggagtcag	27540
	agaaaggaga gggtagaggca gagaccaaag gagggaaaca cttgtgctct acagctactg	27600
	actgagtacc agctgctggt cagacagcca atgccaaggc tcggctgac atggcacctc	27660
30	gtgggactcc tagcccagtg ctggcagagg ggagtgtgta atgggtgcatg gtttgatatt	27720
	gatctgaatg tggccagcc ctagtctcct tccagttgct gggataaagc accctgacca	27780
	aagctacttt tttgtttgtt tgttttgggt tggttttgtt tggtttttcg aggcagggtt	27840
	tctctgtatc accctagctg tcctggaact cactctgtag accaggctgg cctcgaactc	27900
	agaaatcccc ctgcctctgc ctcttaagt ctggaattaa aggcctgcgc caccactgcc	27960
35	ggcccaaagc tactttaaga gagagagagg aatgtataag tattataatt ccaggttata	28020
	gttcattgct gtagaattgg agtcttcata ttccaggtaa tctccacag acatgccaca	28080

	aaacaacctg ttttacgaaa tctctcatgg actcccttcc ccagtaattc taaactgtgt	28140
	caaattctaca agaaatagtg acagtcacag tctctaactg tttgggcatg agtctgaagt	28200
	ctcattgcta agtactggga agatgaaaac tttacctagt gtcagcattt ggagcagagc	28260
	ctttgggatt tgagatggtc ttttgagag ctcctaattg ctacatggag agagggggcc	28320
5	tgggagagac ccatacacct tttgctgcct tatgtcacct gacctgctcc ttgggaagct	28380
	ctagcaagaa ggccctccct ggatcaccca ccaccttgca cctccagaac tcagagccaa	28440
	attaaaacttt ctgtttactg tcttcaaagc acagtcggtc tgggttgat cactgtcaat	28500
	gggaaacaga cttgcctgga tggataactt gtacattgca taatgtctag aaatgaaaag	28560
	tcctatagag aaaaagaaaa ttagctggca cacagataga ggccctggag gaggctggct	28620
10	ttgtccctcc cgaggaggtg gcgagtaagg tgtaaatgtt catggatgta aatgggcccc	28680
	tatatgaggg tctggggtaa caagaaggcc tgtgaatata aagcactgaa ggtatgtcta	28740
	gtctggagaa ggtcactaca gagagtcttc caactcagtg ccatacacaca cacacacaca	28800
	cacacacaca cacacacaca cacacacaca ccacaaagaa aaaaaggaag aaaaatctga	28860
	gagcaagtac agtacttaaa attgtgtgat tgtgtgtgtg actctgatgt cacatgctca	28920
15	tcttgcccta tgagttgaaa accaaatggc ccctgagagg cataacaacc aactgttg	28980
	ctgtgtgttc acgtttttct taaagcgtct gtctggtttg ctgttagcat caggcagact	29040
	tgcagcagac tacatatgct cagccctgaa gtccctctag ggtgcatgtc tcttcagaat	29100
	ttcagaaagt catctgtggc tccaggaccg cctgcactct cctctgccc cgaggctgca	29160
	gactctaggg tgggggtggaa gcaacgctta cctctgggac aagtataaca tgttcgcttt	29220
20	tctttccctc tgtggctcca acctggacat aaaatagatg caagctgtgt aataaatatt	29280
	tcctcccgtc cacttagttc tcaacaataa ctactctgag agcacttatt aatagggtggc	29340
	ttagacataa gctttggctc attccccac tagctcttac tcttttaact ctttcaaact	29400
	attctgtgtc tccacatgg ttagttacct ctccctccat cctgggttcgc tcttccctc	29460
	gagtcgccc cagtgtctct aggtgatgct tgtaagatat tctttctaca aagctgagag	29520
25	tgggtggcact ctgggagttc aaagccagcc tgatctacac agcaagctcc aggatatcca	29580
	gggcaatgtt gggaaaacct ttctcaaaca aaaagagggg ttcagttgtc aggaggagac	29640
	ccatgggtta agaagtctag acgagccatg gtgatgcata cctttcatcc aagcacttag	29700
	gaggcaaaga aaggtgaaac tctttgactt tgaggccagc taggttacat agtgataccc	29760
	tgcttagtgt gtgtgtgtgt gtgtgtgtgt gtgtgtgtgt gtgtgttaatt taaaagtcta	29820
30	aaaatgcatt cttttaaaaa tatgtataag tatttgctg cacatatgta tgtatgtatg	29880
	tataccatgt gtgtgtctgg tgctgaagga ctaggcatag actccctaga actagagtca	29940
	tagacagtgt tgacactccc caacccccca ccatgtgggt gcttgaagct aaactcctgt	30000
	cctttgtaaa gcagcaggtg tctatgaacc ctgaaccatc tctccagtct ccagatgtgc	30060
	attctcaaag aggagtcctt catatttccc taaactgaac atccttatca gtgagcatcc	30120
35	tcgagtcacc aaagctactg caaacctct tagggaacat tcactattca cttctacttg	30180
	gctcatgaaa cttaagtaca cacacacaaa cacacacaca cacacagagt catgcactca	30240

	caaaagcatg catgtacacc attcttatta gactatgctt tgctaaaaga cttctcctaga	30300
	tacttttaaaa catcacttct gcccttttggg gggcagggtc caagattggg actggcgta	30360
	tggaaactga acaaggtaga gatctagaaa tcacagcagg tcagaagggc cagcctgtac	30420
	aagagagagt tccacacctt ccaggaacac tgagcagggg gctgggacct tgcctctcag	30480
5	cccaagaaac tagtgcggtt cctgtatgca tgcctctcag agattccata agatctgctt	30540
	tctgccataa gatctcctgc atccagacaa gcctagggga agttgagagg ctgcctgagt	30600
	ctctcccaca ggccccctt tgcctggcag tattttttta tctggaggag aggaatcagg	30660
	gtgggaatga tcaaatacaa ttatcaagga aaaagtaaaa aacatatata tatatatatt	30720
	aactgatcta gggagctggc tcagcagtta agagttctgg ctgccccctgc ttcagatctt	30780
10	gctttgattc ccagcaccca catgatggct ttcaactgta tctctgcttc caggggatcc	30840
	aacagcctct tctgacctcc atagacaaga cctagtcctc tgcaagagca ccaaatgctc	30900
	ttatctgttg atccatctct ctagcctcat gccagatcat ttaaaactac tggacactgt	30960
	cccattttac gaagatgtca ctgcccagtc atttgccatg agtggatatt tcgattcttt	31020
	ctatgttctc acccttgcaa ttataagaa agatatctgc atttgtctcc tgagagaaca	31080
15	aaggggtggag ggctactgag atggctctag gggtaaaggt gcttgccaca aaatctgaca	31140
	acttaagttt ggtcttggaa tccacatggg ggagagagag aagagattcc cgtaagttgt	31200
	cctcaaaact cccacacatg tgetgtggct tatgtgtaac cccaataagt aaagatagtt	31260
	ttaaacacta cataaggtag ggtttcttca tgaccccaag gaatgatgcc cctgatagag	31320
	cttatgctga aaccccatct ccattgtgcc atctggaaag agacaattgc atcccgga	31380
20	cagaatcttc atgaatggat taatgagcta ttaagaaagt ggcttggtta ttgcacatgc	31440
	tggcggcgta atgacctcca ccatgatgtt atccagcatg aaggctctca ccagaagtca	31500
	tacaaatctt cttaggcttc cagagtcgtg agcaaaaaaa gcacacctct aaataaatta	31560
	actagcctca ggtagttaac caccgaaaat gaaccaagga agttctaata caaaaccact	31620
	tcccttccct gttcaaacca cagtgcctta ttatctaaaa gataaaacttc aagccaagct	31680
25	tttaggttgc cagtatttat gtaacaacaa ggcccgttga cacacatctg taactcctag	31740
	tactgggcct caggggcaga gacaggtgga gccctggagt ttgaattcca ggttctgtga	31800
	gaaactctgt ctgaaaagac aatatggtga gtgaccggg aggatatctg atattgactt	31860
	ctggccaaca cacagccatc tctgcacatc tgtagtgtga agccttttgc actaagttt	31920
	gccagagtca gagtttgcaa gtgtttgtgg actgaatgca cgtgttgctg gtgatctaca	31980
30	aagtcaccct ccttctcaag ctagcagcac tggcttcggc cagctgctca ttcaagcctc	32040
	tttgagagt catcacgggg atgggggagc agggccctc cctagaacac caagcctgtg	32100
	gttgtttatt caggacatta ttgaggcca agatgacaga taactctatc acttggccaa	32160
	cagtcgggtg ttgcggtgtt aggttatctt tgtgtctgca gaaaacagt caacctggac	32220
	aaaagaaata aatgatatca tttttcatct agggcaactag attccgtggg acaaaaggct	32280
35	ccctggggaa cgaggccggg acagcgggc tcttgagtcg ctatttccgt ctgtcaactt	32340
	ctctaattct ttgatttctt cctctgtctt gtttcttctt tcttgctggg gccagtgga	32400

gtctgtgtac tcacagggag gaggggtggca aagccctggg cctctacggg ctgggggaag 32460  
 gggggaagct gtcggcccag tgactttttc ccccttctct tttctttaga aaccagcttc 32520  
 aatttaagat aatgagcttc ctcattcacg tgtgtcact attcataggg acttatccac 32580  
 ccccgccctg tcaatctggc taagtaagac aagtcaaatt taaaagggaa cgtttttcta 32640  
 5 aaaatgtggc tggaccgtgt gccggcacga aaccagggat ggcggtctaa gttacatgct 32700  
 ctctgccagc cccgggtgcct tttcctttcg gaaaggagac ccggaggtaa aacgaagttg 32760  
 ccaacttttg atgatgggtg gcgcccgggtg actctttaaa atgtcatcca tacctgggat 32820  
 agggaaggct cttcagggag tcatctagcc ctcccttcag gaaaagattc cacttccggg 32880  
 ttagttagct tccacctggg ccccttatccg ctgtctctgc ccactagtcc tcatccatcc 32940  
 10 ggtttccgcc ctcatccacc ttgccccctt agttcctaga aagcagcacc gtagctctgg 33000  
 cagggtgggc attgggtcact ccgctaccac tgttaccatg gccaccaagg tgtcatttaa 33060  
 atatgagctc actgagtcct gcgggatggc ttgggtggta atatgcttgc tgcaaaatcg 33120  
 tgagaactgg agttcaattc ccagcacatg gatgtatttc cagcacctgg aaggcagggg 33180  
 gcagagatct taaagctcct ggccagacag cccagcctaa ttagtaatca gtgagagacc 33240  
 15 ctgtctcaag aaacaagatg gaacatcaaa ggtcaacctc ttgtctccac acacacaaat 33300  
 acacacatgc acatacatcc acacacaggc aaacacatgc acacacctga acacctcca 33360  
 caaatacata cataaaaaaa taaatacata cacacatata tacatacacc aacattccct 33420  
 ctcccttagtc tccctggctac gctcttgcac cccccactaa ggcttcaact tcttctatct 33480  
 cttcatcttg actcctctgt actttgcatg ccttttccag caaaggcttt tctttaaatc 33540  
 20 tccgtcattc ataaactccc tctaaatttc tccccctgcc cttttcttct tctctaggga 33600  
 gataaagaca cacactacaa agtcaccgtg ggaccagttt attcacccac ccacctctgc 33660  
 ttctgttcat ccggccagct aagtagtcca acctctctgg tgcgttacc caggacctgg 33720  
 cttcaccaca gctcctccat gctaccagc cctgcaaacc ttcagcctag cctctgggtc 33780  
 tccaaccagc acaggcccag tctggcttct atgtcctaga aatctccttc attctctcca 33840  
 25 tttccctcct gaatctacca ccttctttct ccttctcct gacctctaatt gtcttggtca 33900  
 aacgattaca aggaagccaa tgaaattagc agtttggggt acctcagagt cagcagggga 33960  
 gctgggatga attcacattt ccaggccctt gctttgctcc ccggattctg acaggcagtt 34020  
 ccgaagctga gtccaggaag ctgaatttaa aatcacactc cagctgggtt ctgaggcagc 34080  
 cctaccacat cagctggccc tgactgagct gtgtctgggt ggcagtgggt ctggtgggtg 34140  
 30 tgggtgggtg ggtgggtggg gtgggtgggt tgggtgggtg ggtgggtggg tgtgtgtgtg 34200  
 ttttctgctt ttacaaaact tttctaattc ttatacaaa gacaaatctg cctcatatag 34260  
 gcagaaagat gacttatgcc tatataagat ataaagatga ctttatgcca cttattagca 34320  
 atagttactg tcaaaagtaa ttctatttat acaccttat acatgggtatt gcttttggtg 34380  
 gagactctaa aatccagatt atgtatttaa aaaaaaatc cccagtcctt aaaagggtgaa 34440  
 35 gaatggaccc agatagaagg tcacggcaca agtatggagt cggagtgtgg agtcctgcc 34500  
 atgggtctga cagaagcatc cagagagggt ccaagacaaa tgcctcgct cctaaggaac 34560

	actggcagcc ctgatgaggt accagagatt gctaagtgga ggaatacagg atcagaccca	34520
	tggaggggct taaagcgtga ctgtagcagc cctccgctga ggggctccag gtgggcgccc	34680
	aaggtgctgc agtgggagcc acatgagagg tgatgtcttg gagtacctc gggtaaccatt	34740
	gtttagggag gtggggattt gtggtgtgga gacaggcagc ctcaaggatg cttttcaaca	34800
5	atgggtgatg agttggaact aaaacagggg ccatcacact ggctcccata gctctgggct	34860
	tgccagcttc cacatctgcc cccaccccc tgtctggcac cagctcaagc tctgtgattc	34920
	tacacatcca aaagaggaag agtagcctac tgggcatgcc acctcttctg gaccatcagg	34980
	tgagagtgtg gcaagcccta ggctcctgtc caggatgcag ggctgccaga taggatgtctc	35040
	agctatctcc tgagctggaa ctatttttagg aataaggatt atgcccggcc ggggttgggc	35100
10	agcaccaccag cagcctgtgc ttgcgtaaaa gcaagtgtctg ttgatttate taaaaacaga	35160
	gccgtggacc caccacacagg acaagtatgt atgcactctgt ttcattgtatc tgaaaagcga	35220
	cacaaccatt tttcacatca tggcatcttc ctaaccccca ttcttttttg ttttgttttt	35280
	ttgagacagg gtttctctgt gtagtcctgg ctgtcctgga actcactttg tagaccaggc	35340
	tggcctcgaa ctacagaaatc ctgggattaa aggtgtgtgc caccacgccc ggccctaacc	35400
15	cccatcttta atgggtgatcc agtggttgaa atttcggggc acacacatgt ccattaggga	35460
	ttagctgctg tcttctgagc tacctggtac aatctttatc ccttggggcc tgggctcctg	35520
	atccctgact cgggcccgat caagtccagt tcctggggcc gatcaagtcc agttcctggg	35580
	cccgaacaag tccagtcctt agctcgatta gctcctctg gctccctggc ctgttcttac	35640
	ttacactctt ccccttgctc tggacttgtt gctttcttta ctcaagttgt ctgccacagt	35700
20	ccctaagcca cctctgtaag acaactaaga taatacttcc ctcaagcacg gaaagtcctg	35760
	agtcaccaca cctctggag gtgtgtggac acatgttcat gcgtgtgggt gcgcttacgt	35820
	acgtgtgc	35828

&lt;210&gt; 18

25 &lt;211&gt; 9301

&lt;212&gt; DNA

&lt;213&gt; Homo sapien

&lt;400&gt; 18

30	tagaggagaa gtctttgggg agggtttgct ctgagcacac ccttttccct cctccgggg	60
	ctgagggaaa catgggacca gccctgcccc agcctgtcct cattggctgg catgaagcag	120
	agaggggctt taaaaaggcg accgtgtctc ggctggagac cagagcctgt gctactggaa	180
	ggtggcgtgc cctcctctgg ctggtaccat gcagctccca ctggccctgt gtctcgtctg	240
	cctgctggta cacacagcct tccgtgtagt ggagggccag ggggtggcagg cgttcaagaa	300
35	tgatgccacg gaaatcatcc ccgagctcgg agagtacccc gagcctccac cggagctgga	360
	gaacaacaag accatgaacc gggcggagaa cggagggcgg cctccccacc acccctttga	420

	gaccaaaggt atgggggtgga ggagagaatt cttagtaaaa gatcctgggg aggtttttaga	450
	aactttctctt tgggaggcctt ggaagactgg ggtagaccca gtgaagattg ctggcctctg	540
	ccagcactgg tcgaggaaca gtcttgccctg gaggtggggg aagaatggct cgctgggtgca	600
	gccttcaaat tcaggtgcag aggcctgagg caacagacgc tggtagagagc ccagggcagg	660
5	gaggacgctg ggggtggtgag ggtatggcat cagggcatca gaacaggctc aggggctcag	720
	aaaagaaaag gtctcaaaaga atctcctcct gggaatatag gagccacgct cagctgctgg	780
	taccactggg aaggggaacaa ggtaagggag cctcccatcc acagaacagc acctgtgggg	840
	caccggacac tctatgctgg tgggtggctgt cccaccaca cagaccaca tcatggaatc	900
	cccaggaggt gaacccccag ctccaagggg aagaaacagg ttccaggcac tcagtaactt	960
10	ggtagtgaga agagctgagg tgtgaacctg gtttgatcca actgcaagat agccctgggtg	1020
	tgtggggggg tgtgggggac agatctccac aaagcagtggt ggaggaaggc cagagaggca	1080
	cccttgcatg gtgcattgcc catggcctgc ccaggagctt ggcacttgaa ggaatgggag	1140
	ttttcggcac agtttttagcc cctgacatgg gtgcagctga gtccaggccc tggaggggag	1200
	agcagcatcc tctgtgcagg agtagggaca tctgtcctca gcagccacc cagtcccaac	1260
15	cttgccctcat tccaggggag ggagaaggaa gaggaacctt gggttcctgg tcaggcctgc	1320
	acagagaagc ccaggtgaca gtgtgcatct ggctctataa ttggcaggaa tcctgaggcc	1380
	atgggggctg ctgaaatgac acttcagact aagagcttcc ctgtcctctg gccattatcc	1440
	aggtggcaga gaagtccact gcccaggctc ctggacccca gccctccccg cctcacaacc	1500
	tgttgggact atgggggtgt aaaaagggca actgcatggg aggccagcca ggacctccg	1560
20	tcttcaaaat ggaggacaag ggcgccctcc cccacagctc cccttctagg caaggctcagc	1620
	tgggctccag cgactgcctg aagggtgtga aggaacccaa acacaaaatg tccacctgct	1680
	tggactccca cgagaggcca cagccctga ggaagccaca tgctcaaaac aaagtcatga	1740
	tctgcagagg aagtgcctgg cctaggggct ctattctcga aaagccgcaa aatgccccct	1800
	tccctgggca aatgcccccc tgaccacaca cacattccag cctgcagag gtgaggatgc	1860
25	aaaccagccc acagaccaga aagcagcccc agacgatggc agtggccaca tctcccctgc	1920
	tgtgcttgtt cttcagagtg ggggtggggg gtggccttct ctgtccctc tctggtttgg	1980
	tcttaagact atttttcatt ctttcttgtc acattggaac tatccccatg aaacctttgg	2040
	gggtggactg gtactcacac gacgaccagc tatttaaaaa gctcccacc atctaagtc	2100
	accataggag acatggtcaa ggtgtgtgca ggggatcagg ccaggcctcg gagcccaatc	2160
30	tctgcctgcc caggagatg caccatgagg cgccattca gataacacag aacaagaaat	2220
	gtgcccagca gagagccagg tcaatgtttg tggcagctga acctgtaggt tttgggtcag	2280
	agctcagggc ccctatggta ggaaagtaac gacagtaaaa agcagccctc agctccatcc	2340
	cccagcccag cctcccatgg atgctcgaac gcagagcctc cactcttgcc ggagccaaaa	2400
	ggtgctggga cccaggggaa gtggagtccg gagatgcagc ccagcctttt gggcaagtcc	2460
35	ttttctctgg ctgggcctca gtattctcat tgataatgag ggggttgagc aactgcctt	2520
	tgattccttt caagtctaata gaattcctgt cctgatcacc tccccctcag tccctgcct	2580



	ccacagcagc	tgccttgatt	tattaccttc	aattaacctc	tactcctttc	tccatccctt	2640
	gtccacccct	cccaagtggc	tggaaaagga	atttgggaga	agccagagcc	aggcagaagg	2700
	tgtgctgagt	acttaccctg	cccaggccag	ggacctgcg	gcacaagtgt	ggcttaaate	2760
	ataagaagac	cccagaagag	aaatgataat	aataatacat	aacagccgac	gctttcagct	2820
5	atatgtgcca	aatggatatt	tctgcattgc	gtgtgtaatg	gattaactcg	caatgcttgg	2880
	ggcggcccat	tttgagaca	ggaagaagag	agagggttaag	gaacttgccc	aagatgacac	2940
	ctgcagttag	cgatggagcc	ctgggtgttg	aacccagca	gtcatttggc	tccgagggga	3000
	caggggtgcg	aggagagctt	tccaccagct	ctagagcatc	tgggaccttc	ctgcaataga	3060
	tgttcagggg	caaaagcctc	tggagacagg	cttggcaaaa	gcagggctgg	ggtggagaga	3120
10	gacgggcccg	tccagggcag	gggtggccag	gcgggcccgc	accttcacgc	gcgcctctct	3180
	ccacagacgt	gtccgagtac	agctgcccg	agctgcactt	caccgcctac	gtgaccgatg	3240
	ggccgtgccg	cagcgcgaag	ccggtcaccc	agctgggtgtg	ctccggccag	tgcggcccgg	3300
	cgcgcctgct	gcccacgcc	atcggcccg	gcaagtgggtg	gcgacctagt	gggcccgaact	3360
	tccgctgcat	ccccgaccgc	taccgcgcgc	agcgcgtgca	gctgctgtgt	cccgggtggtg	3420
15	aggcgcgcgc	cgcgcgcaag	gtgcgcctgg	tggcctcgtg	caagtgcaag	cgccctcacc	3480
	gcttccacaa	ccagtcggag	ctcaaggact	tggggaccga	ggccgctcgg	ccgcagaagg	3540
	gccggaagcc	gcggccccgc	gcccggagcg	ccaaagccaa	ccaggccgag	ctggagaacg	3600
	cctactagag	cccgcgcgcg	ccccccccca	ccggcgggcg	ccccggccct	gaaccgcgcg	3660
	cccacatttc	tgtcctctgc	gcgtgggttg	attgtttata	tttcattgta	aatgcctgca	3720
20	accaggggca	gggggctgag	accttccagg	ccctgaggaa	tcccgggccc	cggcaaggcc	3780
	ccccctagcc	cgcagctga	ggggctccac	ggggcagggg	agggaattga	gagtcacaga	3840
	cactgagcca	cgcagccccg	cctctggggc	cgcctacctt	tgctgggtccc	acttcagagg	3900
	aggcagaaat	ggaagcattt	tcaccgccct	ggggttttta	gggagcgggtg	tgggagtggg	3960
	aaagtccagg	gactgggtta	gaaagtggga	taagattccc	ccttgccact	cgctgcccct	4020
25	cagaaagcct	gaggcgtgcc	cagagcacia	gactgggggc	aactgtagat	gtgggtttcta	4080
	gtcctggctc	tgccactaac	ttgctgtgta	accttgaact	acacaattct	ccttcggggac	4140
	ctcaatttcc	actttgtaaa	atgaggggtg	aggtgggaat	aggatctcga	ggagactatt	4200
	ggcatatgat	tccaaggact	ccagtgcctt	ttgaatgggc	agaggtgaga	gagagagaga	4260
	gaaagagaga	gaatgaatgc	agttgcattg	attcagtgcc	aaggctcactt	ccagaattca	4320
30	gagttgtgat	gctctcttct	gacagccaaa	gatgaaaaac	aaacagaaaa	aaaaaagtaa	4380
	agagtctatt	tatggctgac	atattttacg	ctgacaaact	cctggaagaa	gctatgctgc	4440
	ttcccagcct	ggcttccccg	gatgtttggc	tacctccacc	cctccatctc	aaagaaataa	4500
	catcatccat	tggggtagaa	aaggagaggg	tccgaggggtg	gtgggagggga	tagaaatcac	4560
	atccgccccca	acttccccaa	gagcagcatc	cctcccccca	cccatagcca	tgttttaaaag	4620
35	tcaccttccg	aagagaagtg	aaaggttcaa	ggacactggc	cttgcaggcc	cgagggagca	4680
	gccatcacaa	actcacagac	cagcacatcc	cttttgagac	accgccttct	gcccaccact	4740

	cacggacaca tttctgccta gaaaacagct tcttactgct cttacatgtg atggcatatc	4800
	ttacactaaa agaattattat tgggggaaaa actacaagtg ctgtacatat gctgagaaac	4860
	tgcagagcat aatagctgcc acccaaaaat ctttttgaaa atcatttcca gacaacctct	4920
	tactttctgt gtagtcttta attgttaaaa aaaaaaagtt ttaaacagaa gcacatgaca	4980
5	tatgaaagcc tgcaggactg gtcgtttttt tggcaattct tccacgtggg acttgtccac	5040
	aagaatgaaa gtagtggttt ttaaagagtt aagttacata tttattttct cacttaagtt	5100
	atztatgcaa aagttttttt tgtagagaat gacaatgtta atattgcttt atgaattaac	5160
	agtctgttct tccagagtcc agagacattg ttaataaaga caatgaatca tgaccgaaag	5220
	gatgtggtct cattttgtca accacacatg acgtcatttc tgtcaaagtt gacaccttc	5280
10	tcttggtcac tagagctcca accttggaca cacctttgac tgcctctctgg tggcctctgt	5340
	ggcaatttatg tcttcttttg aaaagtcattg tttatccctt cctttccaaa cccagaccgc	5400
	atcttctcac ccaggggcatg gtaataacct cagccttgta tcttttttagc agcctccctt	5460
	ccatgctggc tcccaaaaatg ctgtttctcat tgtatcactc ccttgcctcaa aagccttcca	5520
	tagctccccc ttgcccagga tcaagtgcag tttccctatc tgacatggga ggccttctct	5580
15	gcttgactcc cactccccc tccaccaagc ttcctactga ctccaaatgg tcatgcagat	5640
	ccctgcttcc ttagtttgcc atccacactt agcaccccca ataactaatc ctctttcttt	5700
	aggattcaca ttacttgrca tctcttcccc taaccttcca gagatgttcc aatctcccat	5760
	gateccctctc tctcttgagg ttccagcccc ttttgtctac accactactt tggttccctaa	5820
	ttctgttttc catttgacag tcattcatgg aggaccagcc tggccaagtc ctgcttagta	5880
20	ctggcataga caacacaaag ccaagtacaa ttcaggacca gctcacagga aacttcatct	5940
	tcttcgaagt gtggatttga tgcctcctgg gtagaaatgt aggatcttca aaagtgggcc	6000
	agcctcctgc acttctctca aagtctcgcc tccccagggt gtcttaatag tgcctggatgc	6060
	tagctgagtt agcatcttca gatgaagagt aaccctaaag ttactcttca gttgccctaa	6120
	ggtgggatgg tcaactggaa agctttaaat taagtccagc ctaccttggg ggaacccacc	6180
25	cccacaaaga aagctgaggt cctcctgat gacttgtcag ttttaactacc aataaccac	6240
	ttgaattaat catcatcatc aagtctttga taggtgtgag tgggtatcag tggccgggtcc	6300
	cttcctgggg ctccagcccc cgaggaggcc tcagttagcc cctgcagaaa atccatgcat	6360
	catgagtgtc tcagggccca gaatatgaga gcaggtagga aacagagaca tcttccatcc	6420
	ctgagaggca gtgcgggtcca gtgggtgggg acacgggctc tgggtcaggt ttgtgttgtt	6480
30	tgtttgtttg ttttgagaca gagtctcgct ctattgcccc ggctggagtg cagtgtcaca	6540
	atctcggctt actgcaactt ctgccttccc ggattcaagt gattctcctg cctcagcctc	6600
	cagagtagct gggattacag gtgcgtgcca ccacgcctgg ctaatttttg tatttttgat	6660
	agagacgggg tttcaccatg ttggccaggc tagtctcgaa ctcttgacct caagtgatct	6720
	gcctgcctcg gcctcccaaa gtgctgggat tacaggcgtg agccaccaca cccagcccca	6780
35	ggttgggtgtt tgaatctgag gagactgaag caccaagggg ttaaattgtt tgcccacagc	6840
	catacttggg ctgagttcct tgccttacct ctacttgag ctgcttagaa cctgggtgggc	6900

	acatggggcaa taaccagggtc acactgtttt gtaccaagtg ttatgggaat ccaagatagg	6960
	agtaatttgc tctgtggagg ggatgagggg tagtgggttag ggaaagcttc acaaagtggg	7020
	tgttgcttag agattttcca ggtggagaag ggggtttcta ggcagaaggc atagcccaag	7080
	caaagactgc aagtgcattg ctgctcatgg gtagaagaga atccaccatt cctcaacatg	7140
5	taccgagtc tggccatgtg caaggcaaca tgggggtacc aggaattcca agcaatgtcc	7200
	aaacctaggg tctgttttct gggacctgaa gatacaggat ggatcagccc aggctgcaat	7260
	cccattacca cgaggggggaa aaaaacctga aggttaaat gtaggtcggg ttagaggtta	7320
	tttatggaaa gttatattct acctacatgg ggtctataag cctggcgcca atcagaaaag	7380
	gaacaaacaa cagacctagc tgggaggggc agcatcttgt tgtagggggc ggggcacatg	7440
10	ttctgggggt acagccagac tcagggtttg tattaatagt ctgagagtaa gacagacaga	7500
	gggtagaag gaaataggtc cctttctctc tctctctctc tctctctctc actctctctc	7560
	tctctcacac acacacacag acacacacac acgctctgta ggggtctact tatgctccaa	7620
	gtacaaatca ggccacattt acacaaggag gtaaaggaaa agaacgttgg aggagccaca	7680
	ggaccccaaa attccctgtt ttcttgaat caggcaggac ttacgcagct gggaggggtg	7740
15	agagcctgca gaagccacct gcgagtaagc caagttcaga gtcacagaca ccaaaagctg	7800
	gtgccatgtc ccacaccgc ccacctcca cctgctcctt gacacagccc tgtgctccac	7860
	aacctggctc ccagatcatt gattatagct ctggggcctg caccgtcctt cctgccacat	7920
	ccccacccca ttcttggaa ctgccctctg tcttctccct tgtccaaggg caggcaaggg	7980
	ctcagctatt gggcagcttt gaccaacagc tgaggctcct ttgtggctg gagatgcagg	8040
20	aggcagggga atattcctct tagtcaatgc gaccatgtgc ctggtttgcc cagggtggctc	8100
	tcttttacac ctgtaggcca agcgttaatta ttaacagctc ccacttctac tctaaaaaat	8160
	gacccaatct gggcagtaaa ttatatggtg cccatgctat taagagctgc aacttgctgg	8220
	gcgtgggtggc tcacacctgt aatcccagta ctttgggacg tcaaggcggg tggatcacct	8280
	gaggtcacga gttagagact ggcctggcca gcatggcaaa accccatctt tactaaaaat	8340
25	acaaaaatta gcaaggcatg gtggcatgca cctgtaatcc caggtaactg ggaggctgag	8400
	acaggagaat ggcttgaacc caggaggcag aggttgagct gagccaagat tgtgccactg	8460
	ccctccagcc ctggcaacag agcaagactt catctcaaaa gaaaaaggat actgtcaatc	8520
	actgcaggaa gaacctcagg aatgaatgag gagaagagag gggctgagtc accatagtgg	8580
	cagcaccgac tcctgcagga aaggcgagac actgggtcat gggtagtgaa gggtagcctg	8640
30	aatgacgttc tgccttagag accgaacctg agccctgaaa gtgcatgcct gttcatgggt	8700
	gagagactaa attcatcatt ccttggcagg tactgaatcc tttcttacgg ctgccctcca	8760
	atgcccaatt tccctacaat tgtctggggg gcctaagctt ctgcccacca agagggccag	8820
	agctggcagc gagcagctgc aggtaggaga gataggtagc cataaggag gtgggaaaga	8880
	gagatggaag gagaggggtg cagagcacac acctccctg cctgacaact tcctgagggc	8940
35	tggatcatgcc agcagattta aggcggaggc aggggagatg gggcgggaga ggaagtga	9000
	aaggagaggg tggggatgga gaggaagaga gggtagatcat tcattcattc cattgctact	9060

gactggatgc cagctgtgag ccaggcacca cccagctctt gggcatgtgg ttgtaatctt 9120  
 ggagcctcat ggagctcaca gggagtgtctg gcaaggagat ggataatgga cggataacaa 9180  
 ataaacattt agtacaatgt ccgggaatgg aaagttctcg aaagaaaaat aaagctgggtg 9240  
 agcatataga cagccctgaa ggcggccagg ccaggcattt ctgaggaggt ggcatttgag 9300  
 5 c 9301

<210> 19

<211> 21

<212> DNA

10 <213> Artificial Sequence

<220>

<223> Primer for PCR

15 <400> 19

ccggagctgg agaacaacaa g

21

<210> 20

<211> 19

20 <212> DNA

<213> Artificial Sequence

<220>

<223> PRimer for PCR

25

<400> 20

gcactggccg gagcacacc

19

<210> 21

30 <211> 23

<212> DNA

<213> Artificial Sequence

<220>

35 <223> Primer for PCR

<400> 21  
aggccaaccg cgagaagatg acc 23

5 <210> 22  
<211> 21  
<212> DNA  
<213> Artificial Sequence

10 <220>  
<223> Primer for PCR

<400> 22  
gaagtccagg gcgacgtagc a 21

15 <210> 23  
<211> 25  
<212> DNA  
<213> Artificial Sequence

20 <220>  
<223> Primer for PCR

<400> 23  
aagcttggtgta ccatgcagct cccac 25

25 <210> 24  
<211> 50  
<212> DNA  
<213> Artificial Sequence

30 <220>  
<223> Primer for PCR

<400> 24  
35 aagcttctac ttgtcatcgt cgtccttgta gtcgtaggcg ttctccagct 50

<210> 25

<211> 19

<212> DNA

<213> Artificial Sequence

5

<220>

<223> Primer for PCR

<400> 25

10 gcactggccg gagcacacc

19

<210> 26

<211> 39

<212> DNA

15

<213> Artificial Sequence

<220>

<223> Primer for PCR

20

<400> 26

gtcgtcggat ccacgggggtg gcaggcgctc aagaatgat

39

<210> 27

<211> 57

25

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer for PCR

30

<400> 27

gtcgtcaagc ttctacttgt catcgctcctt gtagtcgtag gcgttctcca gctcggc

57

<210> 28

35

<211> 29

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer for PCR

5

<400> 28

gacttggatc ccaggggtgg caggcgctc

29

<210> 29

10

<211> 29

<212> DNA

<213> Artificial Sequence

<220>

15

<223> Primer for PCR

<400> 29

agcataagct tctagtaggc gttctccag

29

20

<210> 30

<211> 29

<212> DNA

<213> Artificial Sequence

25

<220>

<223> Primer for PCR

<400> 30

gacttggatc cgaagggaaa aagaaaggg

29

30

<210> 31

<211> 29

<212> DNA

<213> Artificial Sequence

35

<220>

<223> Primer for PCR

<400> 31

agcataagct tttaatccaa atcgatgga

5

29

<210> 32

<211> 33

<212> DNA

<213> Artificial Sequence

10

<220>

<223> Primer for PCR

<400> 32

15 actacgagct cggccccacc acccatcaac aag

33

<210> 33

<211> 34

<212> DNA

20

<213> Artificial Sequence

<220>

<223> Primer for PCR

25

<400> 33

acttagaagc tttagctct cagccccctc ttcc

34

<210> 34

<211> 66

30

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer for PCR

35

<400> 34



aatctggatc cataacttcg tatagcatatc attatacgaa gttatctgca ggattcgagg 60  
gccccct 66

<210> 35

5

<211> 82

<212> DNA

<213> Artificial Sequence

<220>

10

<223> Primer for PCR

<400> 35

aatctgaatt ccaccggtgt taattaaata acctcgata atgtatgcta tacgaagtta 60  
tagatctaga gtcagcttct ga 82

15

<210> 36

<211> 62

<212> DNA

<213> Artificial Sequence

20

<220>

<223> Primer for PCR

<400> 36

atttaggtga cactatagaa ctcgagcagc tgaagcttaa ccacatggtg gctcacaacc 60  
at 62

<210> 37

<211> 54

30

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer for PCR

35

<400> 37

aacgacggcc agtgaatccg taatcatggc catgctgccg ggtggaggag ggca

54

<210> 38

<211> 31 <212> DNA

5 <213> Artificial Sequence

<220>

<223> Primer for PCR

10 <400> 38

attaccaccg gtgacaccgc ctccctgaca g

31

<210> 39

<211> 61

15 <212> DNA

<213> Artificial Sequence

<220>

<223> Primer for PCR

20

<400> 39

attacttaat taaacatggc gcgccatag gccggcccct aattgcggcg catcgtaat  
t

60

61

25 <210> 40

<211> 34

<212> DNA

<213> Artificial Sequence

30 <220>

<223> Primer for PCR

<400> 40

attacggccg gccgcaaagg aattcaagat ctga

35

34

<210> 41

44

&lt;211&gt; 34

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

5

&lt;220&gt;

&lt;223&gt; Primer for PCR

&lt;400&gt; 41

attacgggcgc gcccttcaca ggccgcaccc agct

34

# INTERNATIONAL SEARCH REPORT

International Application No.

PCT/US 99/27990

## A. CLASSIFICATION OF SUBJECT MATTER

IPC 7 C12N15/12 C07K14/51 C07K14/495 C12N15/63 C12N5/10  
 C07K16/22 C1201/68 C12N15/62 A61K38/18 A61P19/10  
 G01N33/53 A01K67/027

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 7 C07K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	BIRREN ET AL.: "Homo sapiens chromosome 17, clone HRPC905N1, complete sequence" EMBL SEQUENCE DATABASE, 14 November 1997 (1997-11-14), XP002133385 HEIDELBERG DE Ac AC003098 the whole document	1,2, 27-30
X	HILLIER ET AL.: "WshU-Merck EST Project 1997" EMBL SEQUENCE DATABASE, 19 May 1997 (1997-05-19), XP002133386 HEIDELBERG DE Ac AA393939 the whole document	1,2, 27-30
	--- -/-	

☒ Further documents are listed in the continuation of box C.

☒ Patent family members are listed in annex.

\* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search

20 March 2000

Date of mailing of the international search report

07/04/2000

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2  
 NL - 2280 HV Rijswijk  
 Tel. (+31-70) 340-2040. Tx. 31 651 epo nl.  
 Fax: (+31-70) 340-3016

Authorized officer

Ceder, O

# INTERNATIONAL SEARCH REPORT

Int. onal Application No  
PCT/US 99/27990

## C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No
X	<p>BONALDO ET AL.: "Normalization and subtraction: two approaches to facilitate gene discovery" EMBL SEQUENCE DATABASE, 4 September 1998 (1998-09-04), XP002133484 HEIDELBERG DE Ac AI113131 the whole document &amp; BONALDO ET AL.: "Normalization and subtraction: two approaches to facilitate gene discovery" GENOME RES, vol. 6, no. 9, 1996, pages 791-806,</p>	1,27-30
A	<p>US 5 780 263 A (ADAMS MARK D ET AL) 14 July 1998 (1998-07-14) cited in the application column 1, line 11 - line 13 column 1, line 40 - line 42 column 1, line 66 -column 2, line 47 column 9, line 50 - line 53 column 11, line 15 - line 37</p>	1-22,32, 61-67, 73-79
A	<p>US 5 453 492 A (BUETZOW RALF ET AL) 26 September 1995 (1995-09-26)</p> <p>abstract column 3, line 60 -column 8, line 30</p>	1-3,8,9, 11-13, 15-22, 59,61-67
A	<p>WO 91 13152 A (LUDWIG INST CANCER RES) 5 September 1991 (1991-09-05)</p> <p>the whole document</p>	1-3,8, 11,13, 15,17,32
A	<p>HSU D R ET AL: "The Xenopus dorsalizing factor Gremlin identifies a novel family of secretes proteins that antagonize BMP activities" MOLECULAR CELL,US,CELL PRESS, CAMBRIDGE, MA, vol. 1, no. 5, April 1998 (1998-04), pages 673-683, XP002113640 ISSN: 1097-2765 cited in the application abstract page 676, left-hand column, line 10 - line 14</p>	17
A	<p>WO 92 06693 A (FOX CHASE CANCER CENTER) 30 April 1992 (1992-04-30)</p>	

# INTERNATIONAL SEARCH REPORT

International application No.

PCT/US 99/ 27990

## Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☒ Claims Nos.: 57 and 58  
because they relate to subject matter not required to be searched by this Authority, namely:  
see PCT/ISA/210
2. ☐ Claims Nos.:  
because they relate to parts of the International Application that do not comply with the prescribed requirements to such  
an extent that no meaningful International Search can be carried out, specifically:
3. ☐ Claims Nos.:  
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

## Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. ☐ As all required additional search fees were timely paid by the applicant, this International Search Report covers all  
searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment  
of any additional fee.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this International Search Report  
covers only those claims for which fees were paid, specifically claims Nos.:
4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is  
restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest.
- ☐ No protest accompanied the payment of additional search fees.

# INTERNATIONAL SEARCH REPORT

International Application No. PCT/US 99 27990

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

Continuation of Box I.1

Although claims 57 and 58 are directed to a method of treatment of the human/animal body, the search has been carried out and based on the alleged effects of the compound/composition.

-----

Continuation of Box I.1

Rule 39.1(iv) PCT - Method for treatment of the human or animal body by therapy

# INTERNATIONAL SEARCH REPORT

information on patent family members

International Application No

PCT/US 99/27990

Patent document cited in search report		Publication date	Patent family member(s)	Publication date
US 5780263	A	14-07-1998	CA 2220912 A WO 9639486 A AU 2766595 A EP 0871705 A JP 11506918 T	12-12-1996 12-12-1996 24-12-1996 21-10-1996 22-06-1999
US 5453492	A	26-09-1995	NONE	
WO 9113152	A	05-09-1991	US 5177197 A AU 649026 B AU 7449591 A CA 2076979 A DE 69131572 D DE 69131572 T EP 0517779 A JP 5504888 T	05-01-1993 12-05-1994 18-09-1991 28-08-1991 07-10-1999 23-12-1999 16-12-1992 29-07-1993
WO 9206693	A	30-04-1992	AU 662304 B AU 8957591 A CA 2094608 A EP 0554376 A JP 6502311 T	31-08-1995 20-05-1992 23-04-1992 11-08-1993 17-03-1994